U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

MSc

2.º
CICLO

FCUP
IPATIMUP
2013

U.PORTO

A study on Pharmacogenetic polymorphisms in the Portuguese Gypsies

Joana Maria Santiago Teixeira

FC

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

IPATIMUP
Instituto de Patologia e Imunologia Molecular da Universidade do Porto

MSc

# A study on Pharmacogenetic polymorphisms in the Portuguese Gypsies

Joana Maria Santiago Teixeira

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto e Instituto de Patologia e Imunologia Molecular da Universidade do Porto em Genética Forense

2013

# A study on Pharmacogenetic polymorphisms in the Portuguese Gypsies

Joana Maria Santiago Teixeira
Mestrado de Genética Forense
Departamento de Biologia
2013

**Orientador**
Doutora Maria Sofia Pacheco Quental, Investigadora Doutorada do Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP)

**Coorientador**
Professora Doutora Maria João Prata Martins Ribeiro, Professora Associada com Agregação na Faculdade de Ciências da Universidade do Porto e Investigadora no Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP)

IPATIMUP

Instituto de Patologia e Imunologia Molecular da Universidade do Porto

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Dissertação de candidatura ao grau de Mestre em Genética Forense submetida à Faculdade de Ciências da Universidade do Porto.

O presente trabalho foi desenvolvido sob a orientação científica da Doutora Maria Sofia Pacheco Quental e foi realizado no Instituto de Patologia e Imunologia Molecular da Universidade do Porto.

Dissertation for applying to a Master's Degree in Forensic Genetics, submitted to the Faculty of Sciences of the University of Porto.

The present work was developed under the scientific supervision of Doctor Maria Sofia Pacheco Quental and was done at the Institute of Molecular Pathology and Immunology of University of Porto.

# Acknowledgments

Throughout the realization of my thesis, several persons gave me the support I needed, making possible its successful conclusion. My sincere gratitude is expressed to all of them.

I would like to especially thank my supervisors, Doctor Sofia Quental and Professor Maria João Prata, for their availability, help and support provided during the course of my work. Besides them, I would also like to thank Professor António Amorim for the opportunity to integrate his research group and all his teachings during these two years; finally, I would like to thank my family and friends for all the love and patience demonstrated during all this time.

# Resumo

A Farmacogenética dedica-se ao estudo da influência das variações genéticas humanas na resposta às drogas, e ao longo do tempo tem vindo a adquirir uma importância crescente na prática clínica, com a promessa de tratamentos mais personalizados, que maximizem a eficiência da droga e minimizem a ocorrência de reações de toxicidade. Muitos genes têm sido estudados no campo da farmacogenética, desde genes codificadores de enzimas metabolizadoras, alvos e transportadores das drogas, a fim de compreender o seu papel na variação da resposta às drogas.

Embora muitas populações já tenham sido investigadas em estudos farmacogenéticos, existem ainda populações pouco estudadas neste contexto, como são exemplo os Ciganos da Europa. Devido à sua história e estilo de vida, vivendo em pequenos grupos endogâmicos, os Ciganos têm peculiaridades genéticas únicas, que os tornam em interessantes casos de estudo.

O objetivo principal deste trabalho consistiu em caracterizar uma amostra de Ciganos Portugueses para uma serie de *SNPs* selecionados, localizados em genes com importância farmacogenética e comparar o perfil farmacogenético obtido nesta população, com o perfil da população hospedeira Portuguesa. Neste sentido, uma amostra de 116 Ciganos Portugueses foi genotipada recorrendo ao sistema de *PCR multiplex*, desenhado e otimizado para detetar 10 polimorfismos, localizados em 5 genes, conhecidos por influenciar a resposta as drogas: *CYP2C9* e *CYP2C19* que codificam enzimas com importância na fase I do metabolismo das drogas; *TPMT* e *NAT2* que codificam enzimas intervenientes na fase II do metabolismo e *VKORC1* que codifica um alvo das drogas. Além desta população foi também genotipado um grupo controlo de 70 Portugueses não Ciganos para realizar análises comparativas. Os resultados obtidos para estas duas populações foram comparados com dados anteriormente publicados para várias populações do mundo a fim de avaliar o padrão global de distribuição das frequências.

Para as variações selecionadas não foram detetados desvios ao equilíbrio de Hardy-Weinberg tanto na população Portuguesa Cigana como na População Portuguesa não Cigana e, comparando os resultados das duas populações não foram detetadas diferenças significativas. Contudo, como apenas um pequeno grupo de *SNPs* foi analisado, mais estudos são necessários para explorar a existência de variações adicionais que possam ter um impacto substancial na resposta a drogas nos

indivíduos Ciganos, embora estas sejam irrelevantes (por serem ausentes ou raras) noutros Portugueses.

O presente estudo contribuiu assim para enriquecer a caracterização dos grupos ciganos, abrindo portas para uma melhor compreensão destes grupos populacionais no contexto farmacogenético.

Palavras-Chave: Farmacogenética; População Cigana; Metabolismo das drogas; Alvo das drogas.

# Abstract

Pharmacogenetics is dedicated to the study of the influence of human genetic variations in the drug response and along the time has been acquired an increasing importance in the clinical practice with the promise of more personalised drug treatments that maximises the drug efficacy and minimises the drug toxicity. Many genes have been studied in the field of pharmacogenetics, since drug metabolising enzymes to drug targets and drug transporters, in order to understand their role in variation of drug response.

Although many populations have already been investigated in pharmacogenetic studies, there are still populations scarcely studied in this context, such as the Gypsies from Europe. Due to their history and life style, living in small endogamous groups, the Roma have unique genetic peculiarities, making them an interesting case study.

The main goal of the present work was to characterise a sample of Portuguese Gypsies for a selected battery of SNPs in genes with pharmacogenetic relevance, and to compare the derived pharmacogenetic profile with that of the Portuguese host population. Viewing this, a sample of 116 Portuguese Gypsies was typed with a PCR multiplex system, designed and optimised to detect 10 SNPs located in 5 genes known to influence the response to drugs: *CYP2C9* and *CYP2C19* that encode important enzymes of the phase I of the drug metabolism; *TPMT* and *NAT2* that codify relevant enzymes that intervene in phase II of the drug metabolism and *VKORC1* that codify a drug target. In addition, to perform comparative analyses, a control group of 70 Portuguese non-Gypsy individuals was also genotyped. The results obtained for these two populations were compared with previously published data for several populations around the world, in order to assess global patterns of frequencies distribution.

For the screened variations no departures from Hardy-Weinberg equilibrium were detected in the Portuguese Gypsy and Portuguese non-Gyspy populations. Comparing the results of the two populations no major differences were detected. However, as only a small group of SNPs was analysed, further studies are needed to explore whether additional variations apart from those examined might impact substantially drug response in Gypsies, whilst being irrelevant (due to absence or rarity) in other Portuguese.

The present study has contributed to deepen the characterisation of Gypsy communities, opening new doors to achieve a better knowledge of these population groups in the pharmacogenetic context.

Keywords: Pharmacogenetics; Gypsy population; SNP; Drug metabolism; Drug targets.

# Table of Contents

# Tables Index

# Figures Index

# Abbreviations

| | |
|---|---|
| **µL** | Microlitre |
| **µM** | Micromolar |
| **6-MMP** | 6-methylmercaptopurine |
| **6-MP** | 6-mercaptopurine |
| **6-MTIMP** | 6-methylthioinosine monophosphate |
| **6-TIMP** | 6-thioinosine |
| **6-TG** | 6-thioguanine |
| **6-TU** | 6-thiouric acid |
| **A** | Adenine |
| **ADR** | Adverse Drug Reaction |
| **AZA** | Azathioprine |
| **B.C.** | Before Christ |
| **BLAST** | Basic Local Alignment Search Tool |
| **BLAT** | BLAST-Like Alignment Tool |
| **bp** | base pair |
| **ºC** | Degree Celsius |
| **CYP** | Cytochrome P450 |
| ***CYP1A2*** | Cytochrome P450 1A2 |
| ***CYP2B6*** | Cytochrome P450 2B6 |
| ***CYP2C8*** | Cytochrome P450 2C8 |
| **CYP2C9** | Cytochrome P450 2C9 |
| ***CYP2C18*** | Cytochrome P450 2C18 |
| ***CYP2C19*** | Cytochrome P450 2C19 |
| ***CYP2D6*** | Cytochrome P450 2D6 |
| ***CYP2E1*** | Cytochrome P450 2E1 |
| ***CYP3A4*** | Cytochrome P450 3A4 |
| **ddNTP** | Dideoxynucleotide Triphosphate |
| **DME** | Drug Metabolising Enzyme |
| **DM** | Deficient Methylator |
| **DNA** | Deoxyribonucleic Acid |
| **dNTP** | Deoxynucleotide Triphosphate |
| **EM** | Extensive Metaboliser |
| ***et al*** | et alii |

| | |
|---|---|
| **ExoSAP** | Exonuclease I and Shrimp Alkaline Phosphatase |
| **G** | Guanine |
| **GST** | Glutathione S-transferase |
| **$H^+/K^+$-ATPase** | Proton Pump |
| **He** | Expected Heterozygosity |
| **HM** | High Methylator |
| **Ho** | Observed Heterozygosity |
| ***HPRT*** | Hypoxanthine Phosphoribosyl-Transferase |
| **IA** | Intermediate Acetylator |
| **IM** | Intermediate Metaboliser |
| **IM** | Intermediate Methylator |
| **INH** | Isoniazid |
| **min** | Minute |
| **mRNA** | messenger Ribonucleic Acid |
| **n** | Sample size |
| ***NAT1*** | N-acetyltransferase 1 |
| ***NAT2*** | N-acetyltransferase 2 |
| **OligoCalc** | Oligonucleotide Properties Calculator |
| **PCR** | Polymerase Chain Reaction |
| **PM** | Poor Metaboliser |
| **q** | Frequencies of the minor allele |
| **RA** | Rapid Acetylator |
| **Ref.** | References |
| **rpm** | revolutions per minute |
| **SA** | Slow Acetylator |
| **SAP** | Shrimp Alkaline Phosphatase |
| **SBE** | Single Base Extension primers |
| **sec** | Second |
| **SNP** | Single Nucleotide Polymorphism |
| **TGN** | Thioguanine Nucleotide |
| **Tm** | Melting temperature |
| ***TPMT*** | Thiopurine S-Methyltransferase |
| **UCSC** | University of California, Santa Cruz |
| **UM** | Ultrarapid Metaboliser |
| **VKA** | Vitamin k antagonist |
| ***VKORC1*** | Vitamin K epoxide Reductase Complex subunit 1 |

| | |
|---|---|
| **XME** | Xenobiotic Metabolising Enzyme |
| **XO** | Xanthine oxidase |

# I. INTRODUTION

## 1. Pharmacogenetics

Pharmacogenetics arose from the challenge of understanding how the genetic variations influence the different drug responses that many individuals present. In 510 B.C., Pythagoras had already described that some individuals, after the ingestion of fava beans, manifested unusual reactions, today known as haemolytic anaemia, while others did not showed them (Kupiec *et al.*, 2006; Pelotti and Bini, 2011). However, it was only in 1950 that was clearly documented the first association between inherited genetic information and variable drug response, leading to the acknowledgement of the Pharmacogenetics as a science (Evans and McLeod, 2003; Kalow, 2006; Lanfear and McLeod, 2007).

Nowadays is well established that interindividual variation in drug disposition and effectiveness is in great part due to genetic factors, while non-genetic factors such as age, gender, co-medication or even the nature of the disease may also contribute to variability in drug response (Evans and McLeod, 2003; Pilgrim *et al.*, 2011).

The genetic influence is mainly due to the presence of variations in DNA sequence, known as genetic polymorphisms when the frequency of the minor allele is ≥ 1% in a population. In the human genome, the single nucleotide polymorphisms (SNPs) are the most common source of genetic polymorphisms accounting for approximately 90% of all genetic variations. The SNPs result from changes in single nucleotides of the DNA sequence that sometimes can lead to the alteration of the encoded protein, as well as to the alteration in the quantity of protein expressed, depending on the nucleotide changed and its location (Lanfear and McLeod, 2007; Kudzi *et al.*, 2011; Pilgrim *et al.*, 2011).

When genetic polymorphisms occur in genes that encode enzymes responsible for the absorption, distribution, metabolism or excretion of drugs, such as drug metabolising enzymes, drug transporters and drug targets, that can lead to the alteration of kinetics and dynamics of drugs in the human organism, thus causing variability in drug responses (Evans and McLeod, 2003; Weinshilboum and Wang, 2006; Pilgrim *et al.*, 2011). It is on these processes that Pharmacogenetics concentrates its efforts, trying to understand how genetic polymorphisms affect either pharmacokinetics or pharmacodynamics of drugs. Indeed, given that pharmacokinetics comprehends the pathways that a drug takes through the organism, from absorption, to

metabolism and excretion, whereas pharmacodynamics respect to how drugs interact with receptors to initiate the drug response, it is understandable that both dynamics can determine drug effects and responses (Johnson, 2003; Weinshilboum and Wang, 2004; Weinshilboum and Wang, 2006) (Figure 1).

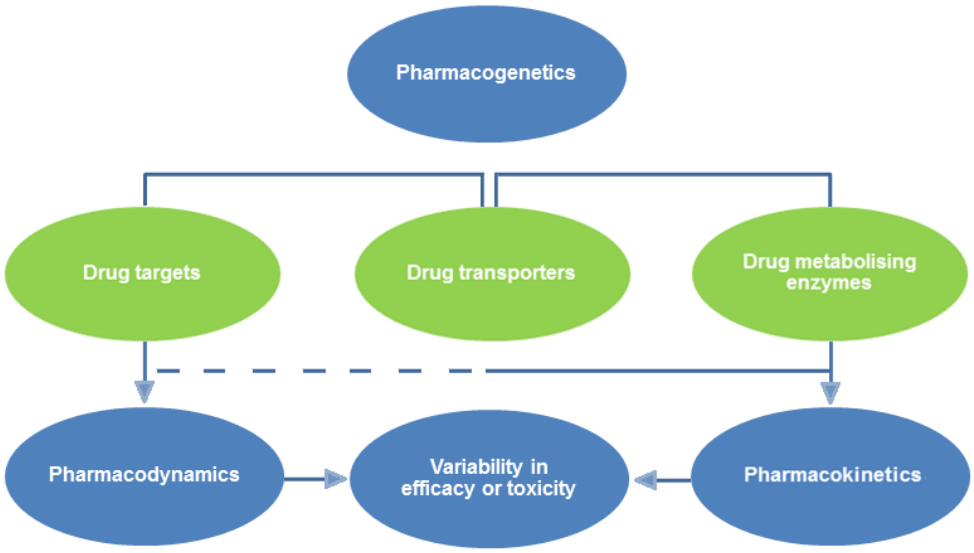


Figure 1. The functional components of Pharmacogenetics. The drug metabolising enzymes and drug transporters frequently influence the pharmacokinetics, while the drug targets commonly contribute for the variability in pharmacodynamics. The dashed line indicates that drug transporters are sometimes the drug targets. Adapted from Johnson, (2003).

The information acquired through pharmacogenetic studies can be a useful tool in several areas such as the clinical practice and the forensic analysis, since it allows to improve the relation between patient and drug, avoiding unexpected results in therapeutic treatments and contributing to the interpretation of deaths that remain unclear (Pilgrim *et al.*, 2011). In the following section these two subjects will be explored in more detail.

## 1.1 Pharmacogenetic applications

As mentioned in the previous section it is frequent to observe that different patients may react differently to the same drug and dose: some develop adverse drug reactions (ADRs), which sometimes can be fatal, others react well or with rather efficacy to the treatment, while others do not even respond to the therapy. This variability of responses of individuals to a standardised treatment is a relevant question in clinical practice (Kudzi *et al.*, 2011). The aim of Pharmacogenetics in the medical context is to identify polymorphisms in genes that can affect the effectiveness of a specific drug, diminishing the number of ADRs and treatment failure, therefore allowing the implementation of a more personalised, safe and effective therapy (Spear *et al.*, 2001; Kalow, 2002; Fishbain *et al.*, 2004; Kudzi *et al.*, 2011). This rational administration of drugs is not only beneficial for patients, but also allows the reduction of healthcare costs (Rai *et al.*, 2009). Furthermore, Pharmacogenetics can also play a key role in the development of new drugs, which can be safer if taking into account the genetic information (Daly, 2010).

The recent developments in Pharmacogenetics knowledge have also given an important contribution to the forensic area given its role in the interpretation of circumstances of deaths. Through Pharmacogenetic studies it is possible to provide information about the metabolic capacity of individuals and potential drug outcomes, that can help the evaluation of post-mortem toxicology results and the determination of the involvement of a drug in the death (Pilgrim *et al.*, 2011). That explains why the recruitment of pharmacogenetic data has allowed the clarification, in collaboration with information from the autopsy, clinical history and investigation of the death scene, of the cause and manner of death (Wong *et al.*, 2003; Kupiec *et al.*, 2006; Musshoff *et al.*, 2010).

In particular situations, in which the circumstances of death are difficult to interpret due to non-existence of evidence of trauma or to ambiguous autopsy results, the pharmacogenetic analysis can be a useful tool for the medico-legal investigations, since genetic information can offer an additional knowledge of a possible genetic mutation that predisposes the individual to risk conditions leading to death. For example, in cases that post-mortem toxicological reports indicate a fatal concentration of a particular drug, it is difficult to understand if the death was accidental or suicide. However, with further information of genetic analysis it can be possible to understand if the high concentration of the drug was due to a genetic defect of the individual, making

the cause of death accidental, or if it was due to a high intake, making the cause of death intentional (Sajantila *et al.*, 2010).

Thus, with the association of pharmacogenetic analysis with post-mortem toxicology it can be possible to have new insights into the cause and circumstances of death of an individual (Musshoff *et al.*, 2010; Sajantila *et al.*, 2010).Current challenges

## 1.2 Current challenges

Since the post-genomic era, that clinical application of the Pharmacogenetics appears as being a promise of a more personalized medicine, through the application of the genetic knowledge in the improvement of the drugs and dosages, making the drug administration more safe and efficient (Swen *et al.*, 2007).

Although over time, the knowledge of the role that the genes play in the drug responses has improved and with it have emerged several promising clinical pharmacogenetic tests, their application in patient care is still scarce. Many reasons have been appointed as the cause of this limited application, such as educational, financial, guidelines, social, legal and ethical barriers (Weinshilboum and Wang, 2004; Weinshilboum and Wang, 2006; Swen *et al.*, 2007; Crews *et al.*, 2012; Moaddeb and Haga, 2013).

For implementation of a pharmacogenetic test in clinical practice it is required a strong scientific evidence that it will improve the clinical outcomes of the patients, otherwise its introduction in medical healthcare would not be beneficial (Weinshilboum and Wang, 2004; Swen *et al.*, 2007; Swen and Guchelaar, 2012; Moaddeb and Haga, 2013).

The costs are other barrier to the integration of the Pharmacogenetics into routine clinical. The cost-effectiveness studies for these tests are an important determinant for their implementation since should demonstrate their clinical utility and validity and also prove capable of provide the return in investment to the payers (Scott, 2011; Swen and Guchelaar, 2012). Many recent studies that evaluate the clinical utility and validity of Pharmacogenetics tests have shown results that provide incentive for reimbursement of genetic testing and investment in implementation strategies (Crews *et al.*, 2012).

The education of all healthcare professionals, specially the physicians, about pharmacogenetic testing, influences strongly their implementation into clinical practice, since they are responsible for their prescription. Many studies have indicated that clinicians in general have not much confidence in pharmacogenetic services because

of insufficient training and knowledge. Due to the lack of these two skills it is difficult to the doctors to understand the clinical utility of the tests, how to get them and how to interpret and apply their results into the patients care (Weinshilboum and Wang, 2004; Woelderink *et al.*, 2005; Scott, 2011; Crews *et al.*, 2012; Swen and Guchelaar, 2012).

# 2. The drug metabolism

During the course of our life the human body is exposed to an uncountable number of xenobiotics – foreign elements to the body. Xenobiotics can be food elements, environmental chemicals or even pharmaceuticals, which can cause diverse damages to the body. When these elements enter the organism, a complex of enzymatic mechanisms performs biochemical transformations that convert the components in other, less harmful, trying to minimise the aggression caused by xenobiotic substances (Liska, 1998; Rushmore and Kong, 2002; Gonzalez and Turkey, 2006; Taniguchi and Guengerich, 2012). The mechanism responsible for the biotransformation of xenobiotics is commonly referred to as drug metabolism or drug biotransformation. An essential part of the drug metabolism is the chain of biochemical reactions that transform the foreign elements in components easier to excrete, preventing its accumulation in the body and possible toxic reactions. These reactions are also capable to alter drugs or its precursors in four different forms: an active drug can be inactivated; an active drug can be altered into an active or toxic metabolite; an inactive prodrug can be converted into an active drug and an unexcretable drug can be metabolised into an excretable metabolite (Taniguchi and Guengerich, 2012).

This mechanism of biotransformation of foreign chemicals involves several enzymes, denominated as drug metabolising enzymes or xenobiotic metabolising enzymes (DMEs or XMEs), which are responsible by a number of different steps in the biochemical reactions (Rushmore and Kong, 2002; Xu *et al.*, 2005).

While many tissues and organs are involved in drug metabolism, the liver is the principal organ of drug biotransformation, expressing, therefore, high levels of DMEs. When oral drugs are administrated, the liver metabolism is so efficient that only a small portion of drug passes to the blood circulation – first pass effect (Stachulski and Lennard, 2000; Taniguchi and Guengerich, 2012)

The drug metabolism consists in two main different phases. Phase I, that includes functionalisation reactions (mainly oxidation, hydroxylation, and hydrolysis) by adding or exposing a polar group, and phase II, in which occur conjugation reactions between a polar group, possibly originated in phase I, and endogenous substrates such as glucuronic acid, forming a water-soluble product easier to eliminate (Stachulski and Lennard, 2000).

Although phase I and II reactions can occur independently of one another, normally they occur sequentially. While these are the main phases of the drug

metabolism, references to a phase III can be found in the bibliography, associated with transport of drugs in the organism (Xu *et al.*, 2005; Olinga *et al.*, 2008).

The rate of biotransformation reactions is not the same in all individuals, since it is frequent to observe interindividual differences in the metabolism due to factors such as age, gender, diet, drinking and smoking habits as well as to genetic profile (Pilgrim *et al.*, 2011; Taniguchi and Guengerich, 2012).

Up to now, about 30 gene families of DMEs were already described (Kudzi *et al.*, 2011; Pilgrim *et al.*, 2011).

## 2.1 Phase I of the drug metabolism

The phase I of the drug metabolism is the first enzymatic defence that acts against foreign elements, including therapeutic drugs (Liska, 1998). Usually, the phase I enzymes catalyse several reactions such as oxidation, reduction, hydrolysis, cyclization and decyclization. These reactions frequently lead to the inactivation of the drug, but in some cases they can also bioactivate some drugs that are inactive, called prodrugs, or may even produce toxic molecules that can cause damage to the body if not posteriorly metabolised by phase II enzymes (Liska, 1998; Gonzalez and Turkey, 2006).

Several DMEs are involved in these processes, like esterases, reductases, alcohol dehydrogenases and mainly the cytochrome P450 (CYP) superfamily (Tiwari *et al.*, 2009; Pilgrim *et al.*, 2011).

### 2.1.1 Cytochrome P450 system

The CYP enzymes are a superfamily of heme-thiolate proteins responsible for the metabolism of a great number of endogenous molecules like steroids, fatty acids, prostaglandins, as well as exogenous agents, such as drugs and other xenobiotics (Danielson, 2002; Zanger *et al.*, 2008). These enzymes are located primarily in the endoplasmic reticulum of liver cells, but also in extra-hepatic tissues such as lung, intestines and brain, and are responsible for about 80% of the phase I reactions (Ingelman-Sundberg *et al.*, 2007; Pilgrim *et al.*, 2011).

The human genome comprises several functional *CYP* genes and some pseudogenes, divided in families and subfamilies according to the degree of similarity of their DNA sequence (Zanger *et al.*, 2008). Each CYP isoenzyme is named according

to specific norms that can be illustrated using the example of CYP2C9: the cytochrome P450 designation is CYP, the family designation is indicated by an Arabic number (CYP2), the subfamily designation is characterised by a capital letter (CYP2C) and the last number represents a particular gene (Poolsup *et al.*, 2000; Ma *et al.*, 2002).

The CYP3A4, CYP2D6, CYP2C9, CYP2C19, CYP2B6 and CYP1A2 enzymes have the most crucial role in the drug metabolism, biotransforming more than 90% of the CYP superfamily substrates. Some associations between drug response and genetic polymorphisms of these enzymes have been described, especially with polymorphisms of *CYP2D6, CYP2C9* and *CYP2C19* genes (Hiratsuka *et al.*, 2006).

The polymorphic nature of these CYP enzymes leads to the emergence of differences in their metabolic activity, that correlates with four different phenotypes: the extensive metaboliser (EM) phenotype, when the individuals carry two functional alleles; poor metaboliser (PM) phenotype, when the individuals inherit two defective or deleted alleles; intermediate metaboliser (IM) phenotype, when the individuals inherit one functional and one defective allele or they inherit two partially defective alleles; and ultrarapid metaboliser (UM) phenotype, when more than two active genes are producing a particular CYP enzyme. While the EM show a normal enzymatic activity, comparatively the PM and the UM phenotypes show a lower and a higher metabolic capacity, respectively. On the other hand, the IMs phenotypes show a metabolic capacity between PMs and EMs (Hiratsuka *et al.*, 2006; Ingelman-Sundberg *et al.*, 2007).

Depending on the type of phenotype that the individuals present, the drug metabolism can be variable, consequently affecting the drug effectiveness. As an example, when a prodrug is administrated to an individual with a PM phenotype, therapeutic failure can occur since his low metabolic capacity does not allow the expected bioactivation of the prodrug into an active drug (Musshoff *et al.*, 2010; Pilgrim *et al.*, 2011).

#### 2.1.1.1 Cytochrome P450 2C9 – CYP2C9

The CYP2C9 is an important isoenzyme of the CYP complex that is mainly expressed in the human liver. It is involved in the metabolism of multiple drugs such as warfarin, tolbutamide and phenytoin as well as some endogenous substrates like arachidonic and linolenic acids (Ingelman-Sundberg *et al.*, 2007; Cavallari *et al.*, 2011). The warfarin, an oral anticoagulant widely used for prevention and treatment of

thromboembolism disorders, is one of the more well studied drugs in Pharmacogenetic fields, due to its narrow therapeutic window (Kwon *et al.*, 2011). This drug exerts its pharmacologic effect by inhibiting the vitamin k epoxide reductase complex subunit 1 (VKORC1), an important element in the formation of active clotting factors. Warfarin is a racemic mixture of S-warfarin and R-warfarin where the S-enantiomer is the more active compound and has a greater therapeutic effect (Saminathan *et al.*, 2010; Lam and Cheung, 2012). The enantiomers of warfarin are metabolised by 2 different pathways. S-warfarin is mostly metabolised by CYP2C9 while R-warfarin is mainly metabolised by CYP1A2 and CYP3A4 (D'Andrea *et al.*, 2008; Limdi and Veenstra, 2008; Saminathan *et al.*, 2010) (Figure 2).

Once the more pharmacologically potent S-form of warfarin is mainly biotransformed by CYP2C9, it is expected that variations on *CYP2C9* gene can contribute greatly for the large interindividual variability in warfarin response and dosage requirement (Rieder *et al.*, 2005; D'Andrea *et al.*, 2008; Limdi and Veenstra, 2008).

The *CYP2C9*, located on chromosome 10q24.2, is a polymorphic gene with many different allelic variants identified to date (http://www.cypalleles.ki.se/cyp2c9.htm) (Cavallari *et al.*, 2011; Pilgrim *et al.*, 2011). The *CYP2C9*1*, often designed as the wild-type allele, encodes for a functional product associated to normal enzymatic activity, whereas *CYP2C9*2* (rs1799853, c.430C>T, p.R144C) and *CYP2C9*3* (rs1057910, c.1075A>C, p.I359L) have been identified as the most common reduction-function variants that have an important role in the warfarin metabolism (Pilgrim *et al.*, 2011; Voora and Ginsburg, 2012). Although both conduct to the same metabolic phenotype, while *CYP2C9*2* may cause a decrease of approximately 30% on enzymatic activity, *CYP2C9*3* may lead to reductions of about 80% (Markatos *et al.*, 2008; Kwon *et al.*, 2011).

In the warfarin metabolism, the reduced activity of CYP2C9 leads to a decrease on clearance of S-warfarin and in the warfarin dose requirements, translating into about 19% of dose reduction per allele for *CYP2C9*2* and approximately 33% of dose reduction per allele for *CYP2C9*3* compared to non-carriers. Furthermore, the diminished activity of CYP2C9 increases the risk of bleeding complications (Voora and Ginsburg, 2012).

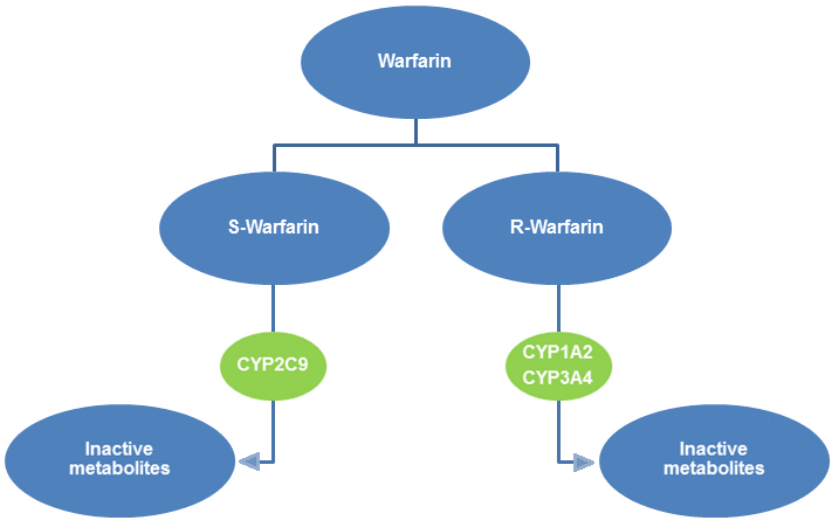


Figure 2. The metabolic pathways of warfarin. The warfarin is a racemic mixture of S- and R-warfarin. While the S-warfarin is mostly metabolised by CYP2C9, the R-warfarin is mainly metabolised by CYP3A4 and CYP1A2.

### 2.1.1.2 Cytochrome P450 2C19 – CYP2C19

The CYP2C19 isoenzyme is part of CYP2C subfamily alongside with the CYP2C8, CYP2C9 and CYP2C18 enzymes, being responsible for the metabolism of various therapeutic drugs such as omeprazole, diazepam, clopidogrel and nelfinavir (Cavallari *et al.*, 2011; Dandara *et al.*, 2011). The clopidogrel, an important antiplatelet agent, is a prodrug that requires hepatic bioactivation in order to produce its active metabolite. This bioactivation occurs in two oxidative steps, where several enzymes participate, including CYP2C19, which contributes for 45% of the first step and for 21% of the second (Figure 3) (Santos *et al.*, 2011). Additionally, CYP2C19 enzyme has also a huge importance in the metabolism of proton pump inhibitors, such as omeprazole. This drug binds to the proton pump ($H^+/K^+$-ATPase) in gastric cells inhibiting the segregation of gastric acid, being a useful drug in the treatment of gastric acid-related disorders (Shirai *et al.*, 2001; Yasui-Furukori *et al.*, 2004; Uno *et al.*, 2007). Omeprazole is metabolised in 2 different pathways, being extensively metabolised by CYP2C19 via (Palovaara *et al.*, 2003; Furuta *et al.*, 2005; Uno *et al.*, 2007) (Figure 4).

The *CYP2C19* gene, located on chromosome 10q24.1-q24.3, is one of the most polymorphic enzymes of the CYP2C members, with many allelic variants described until today (http://www.cypalleles.ki.se/cyp2c19.htm) including *CYP2C19*2* (rs4244285, c.681G>A, splicing defect), *CYP2C19*3* (rs4986893, c.636G>A, p.W212X) and *CYP2C19*17* (rs12248560, c.-806C>T, increased transcription). (Lee *et al.*, 2009;

Tiwari *et al.*, 2009; Daly, 2010). While *CYP2C19*2* and *CYP2C19*3* conduct to PM phenotypes, due to their genetic defect causing an inactive enzyme, *CYP2C19*17* is related to UM phenotype because it is associated with increased levels of gene expression (Brockmöller and Tzvetkov, 2008; Tiwari *et al.*, 2009; Pilgrim *et al.*, 2011).

In case of treatments with clopidogrel, individuals that carry *CYP2C19*2* have low concentration of the active metabolite, which leads to an ineffective inhibition of platelet function and consequently to increased risk of developing cardiovascular events, such as stent thrombosis and myocardial infarction. In contrast, individuals that carry *CYP2C19*17* have a high concentration of the active metabolite due to increased enzymatic activity, which leads to an intensification of the inhibition of platelet function, what on one hand improves the response to antiplatelet treatment with clopidogrel, but on the other hand increases the risk of bleeding events (Santos *et al.*, 2011; Voora and Ginsburg, 2012).

Besides the differences observed in the bioactivation of clopidogrel, alterations in the metabolism of omeprazole and differences in the effectiveness of treatments have also been described. PM phenotype have increased duration of exposure of the proton pumps in parietal cells to high plasma concentration of omeprazole, leading to its inactivation by a longer period, and a better therapeutic efficiency. On the other hand, *CYP2C19*17* is likely to promote increased metabolism of omeprazole that may cause a reduction in acid inhibition and consequently a therapeutic failure (Chaudhry *et al.*, 2008; Hunfeld *et al.*, 2008; Gawrońska-Szklarz *et al.*, 2010).

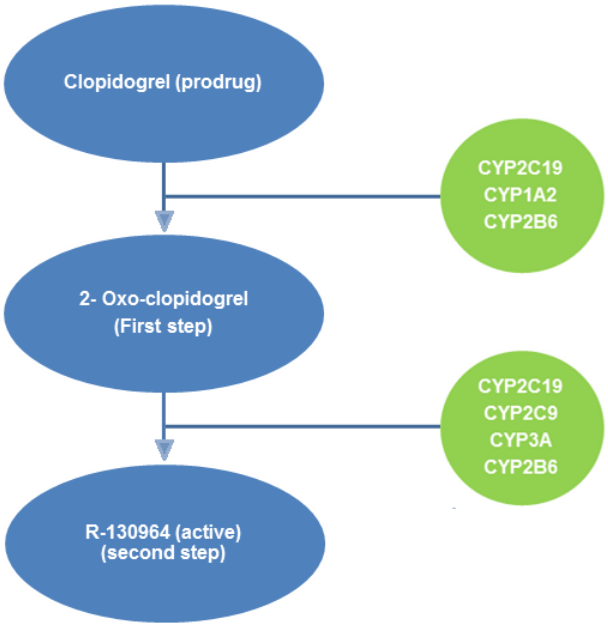


Figure 3. Clopidogrel metabolism. The clopidogrel is a prodrug that requires hepatic bioactivation for its active metabolite exert antiplatelet function. Adapted from Bonello *et al.* (2010) and Holmes *et al.* (2010).

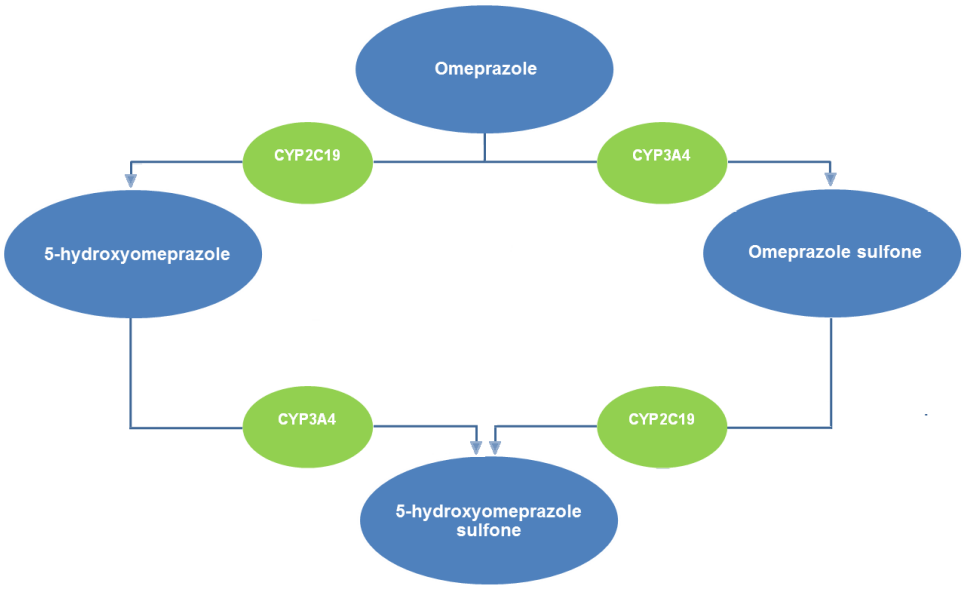


Figure 4. Metabolic pathways of omeprazole by CYP isoenzymes. Omeprazole is mostly metabolised by CYP2C19 into 5-hydroxyomeprazole, and then it is transformed into 5-hydroxyomeprazole sulfone by CYP3A4. A small portion of omeprazole is also metabolised by CYP3A4 into omeprazole sulfone, which is transformed into 5-hydromeprazole sulfone by CYP2C19. Adapted from Furuta *et al.* (2005).

## 2.2 Phase II of the drug metabolism

In the phase II of drug metabolism, generally following the phase I, occurs the binding of an ionised group to the drug, forming a water-soluble product that facilitates their posteriorly extraction through the urine or bile (Liska, 1998; Guo *et al.*, 2011). In this phase occur many conjugation reactions, specifically acetylation, glucuronidation, sulfation and methylation reactions (Pilgrim *et al.*, 2011). In some situations, the phase II metabolite formed can be more pharmacologically active that their parent drug (Stachulski and Lennard, 2000).

The enzymes with the most important role in this transformation are N-acetyltransferase 1 and 2 (NAT1 and NAT2), glutathione transferase, thiopurine S-methyltransferase (TPMT) and uridine diphosphate glucuronosyltransferase (Kudzi *et al.*, 2011).

### 2.2.1 Thiopurine S-methyltransferase – TPMT

The thiopurine S-methyltransferase (TPMT) is a cytosolic enzyme responsible for the S-methylation of thiopurine drugs, such as 6-mercaptopurine (6-MP), 6-thioguanine (6-TG) and azathioprine (AZA) (Kapoor *et al.*, 2009; Pavlovic, 2009). These drugs are mostly used in the treatment of acute leukaemia and chronic inflammatory diseases (Román *et al.*, 2012). When a thiopurine drug is absorbed by the organism, for example AZA, it is transformed in 6-MP in the liver and then is metabolised by 3 different enzymatic pathways mediated by the Hypoxanthine phosphoribosyl-transferase (HPRT), Xanthine oxidase (XO) and TPMT. While the HPRT pathway is the first step to the formation of the active metabolites that exert the cytotoxic effect (6-TGNs), the other two pathways, TPMT and XO, inactivate the drug (Garat *et al.*, 2008; Smith *et al.*, 2009; Zabala-Fernández *et al.*, 2011). The most important inactivation pathway of thiopurines in hematopoietic tissues is the TPMT pathway, since XO is nearly absent in this tissue. Moreover, TPMT activity has been described as inversely correlated to the accumulation of 6-TGNs and hematopoietic toxicity (Krynetski and Evans, 2003) (Figure 5).

The human *TPMT* gene, located in chromosome 6p22.3, exhibits a series of genetic polymorphisms that are responsible for the interindividual and interethnic variability of TPMT activity. Three main TPMT phenotypes can be distinguished: high methylator (HM), when the individuals inherit two functional alleles; intermediate

methylator (IM), when the individuals inherit one functional and one non-functional *TPMT* allele; and deficient methylator (DM), when the individuals inherit two non-functional *TPMT* alleles (Spire-Vayron de la Moureyre *et al.*, 1998; Pavlovic, 2009). In populations of European descent, about 90% of the individuals are HMs, expressing a normal enzymatic activity, approximately 10% are IMs, manifesting an intermediate activity and around 0.3% have DM phenotype, exhibiting a low or undetectable activity of TPMT enzyme (Garat *et al.*, 2008; Zabala-Fernández *et al.*, 2011).

When patients with DM phenotype are treated with a standard dose of thiopurine drugs, they are at high risk of developing severe and even fatal hematopoietic toxicity, since the non-inactivation of the drug by TPMT, results in the excessive accumulation of 6-TGNs in erythrocytes (Spire-Vayron de la Moureyre *et al.*, 1998; Pavlovic, 2009). In these patients the dose of drug administrated must be reduced in order to prevent toxic effects. Patients with intermediate levels of TPMT activity would also benefit from a thiopurine dose adjustment, although the risk of toxic side effects is not so severe as for the DM individuals (Kapoor *et al.*, 2009; Pavlovic, 2009).

The *TPMT*2* (rs1800462, c.238G>C, p.A80P), *TPMT*3A* (rs1800460, c.460G>A, p.A154T and rs1142345, c.719A>G, p.Y240C) *TPMT*3B* (rs1800460, c.460G>A, p.A154T) and *TPMT*3C* (rs1142345, c.719A>G, p.Y240C) are the most common non-functional alleles, usually accounting for 80 to 95% of IMs and DMs phenotypes in Eurasian populations (Garat *et al.*, 2008; Pavlovic, 2009). Many other *TPMT* non-functional alleles have been described, among which is *TPMT*8* (rs56161402, c.644G>A, p.R215H), an allele mainly found in African populations (Oliveira *et al.*, 2007b; Roberts and Barclay, 2012).

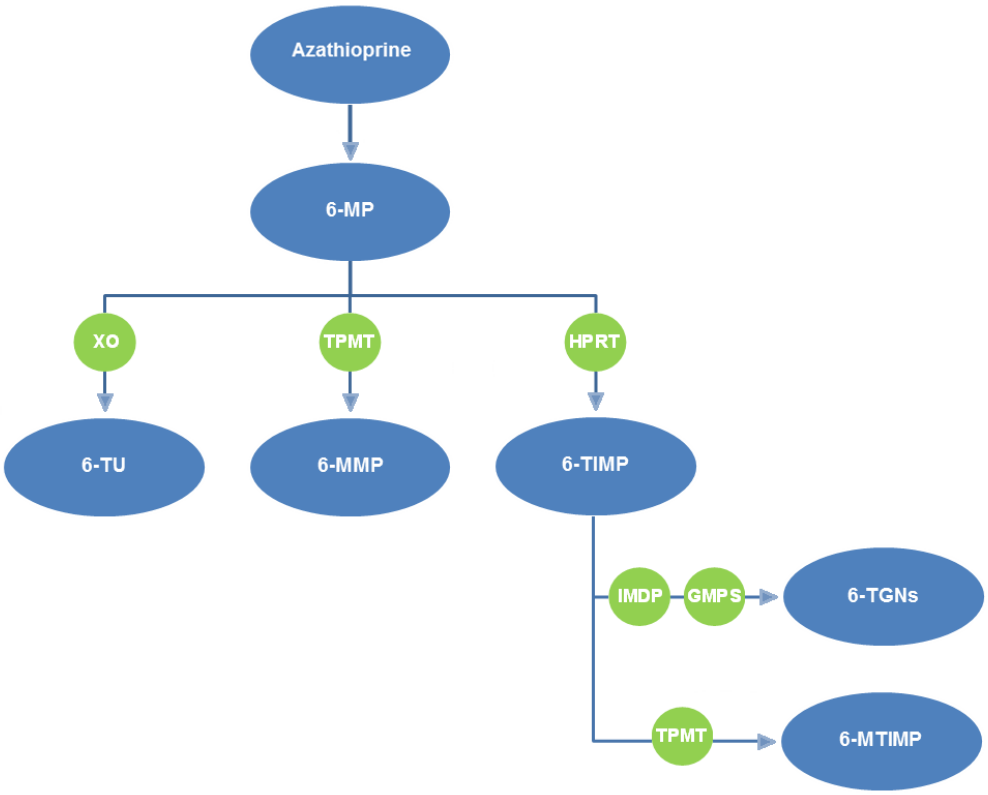


Figure 5. Azathioprine metabolism. This drug is transformed in 6-MP and then is metabolised by the Hypoxanthine phosphoribosyl-transferase (HPRT), Xanthine oxidase (XO) and TPMT pathways. While HPRT pathway is the first step, forming 6-thioinosine monophosphate (6-TIMP) to form the cytotoxic elements 6-TGNs, the remaining two pathways are the initial steps to eliminate thiopurines from the body, forming 6-thiouric acid (6-TU), 6-methylmercaptopurine (6-MMP) and 6-methylthioinosine monophosphate (6-MTIMP). Adapted from Smith *et al.* (2009) and Zabala-Fernández *et al.* (2011).

### 2.2.2 N-acetyltransferase 2 – NAT2

The N-acetyltransferase 2 (NAT2) enzyme intervenes in phase II of drug metabolism and is responsible for the acetylation of various drugs such as isoniazid and sulfamethoxazole as well as several nefarious components like heterocyclic and aromatic amines (Fuselli *et al.*, 2007). The metabolism of isoniazid (INH), an important anti-tuberculosis drug, is one of the most well studied Pharmacogenetic examples. Its metabolism occurs in the liver, where it is biotransformed to acetylisoniazid by NAT2, hydrolysed to acetylhydrazine and posteriorly oxidised by CYP2E1, forming hepatotoxic metabolites. In another metabolic pathway of INH, NAT2 is also responsible for the transformation of hydrazine, the major hepatotoxic metabolite of INH, into acetylhydrazine that subsequently is oxidised by CYP2E1. Furthermore,

NAT2 also participates in the biotransformation of acetylhydrazine into a harmless component, diacetylhydrazine (Teixeira *et al.*, 2011; Lv *et al.*, 2012) (Figure 6).

The *NAT2* gene, located in human chromosome 8p22, is polymorphic with over 30 SNPs described in its coding region, including some rare mutations observed in different human populations. Among these, the most common SNPs are c.191G>A (rs1801279, p.R64Q), c.282C>T (rs1041983, p.Y94Y) c.341T>C (rs1801280,p.I114T), c.481C>T (rs1799929, p.L161L), c.590G>A (rs1799930, p.R197Q), c.803A>G (rs1208, p.K268R), c.857G>A (rs1799931, p.G286E), identified in different populations (Fuselli *et al.*, 2007; Rios-Santos and Magno, 2012; Teixeira *et al.*, 2013).

The combination of *NAT2* polymorphisms is commonly used to infer the NAT2 acetylation status, which are designated as haplotypes (Rios-Santos and Magno, 2012). Three different acetylator phenotypes have been described: rapid (RA), intermediate (IA) and slow acetylator (SA) phenotypes, if the individual inherits 2, 1 or 0 full functioning haplotypes, respectively (Hiratsuka *et al.*, 2006; Fuselli *et al.*, 2007).

Although it is assumed that the phenotype prediction of NAT2 activity through genotyping methods is very useful, so far there is still no agreement about the minimum number of *NAT2* SNPs necessary to infer with accuracy the human acetylator status. The 7-SNP based genotypic approach is the most conventional method used due to its good sensitivity and specificity (Selinski *et al.*, 2011; Rios-Santos and Magno, 2012). However, in a recent study of Selinski *et al*. (2011) the NAT2 phenotypes were determined and the sensibility and specificity of the conventional 7-SNP genotype was compared to that of a recently published tagSNP (rs1495741). When the authors further evaluated whether a novel SNP panel could outperform the two others, they found out that both the tagSNP and the 7-SNP inferred genotypes had a high sensibility to predict slow acetylators, although the tagSNP showed an inferior specificity due to misclassifying some rapid and slow acetylators. Interestingly, a combination of only two SNPs (c.282C>T, rs1041983 and c.341T>C, rs1801280) that had the same sensitivity and specificity as the 7-SNP genotype was also identified. These results, which have been corroborated in the study of Hein *et al.* (2012), may have a practical importance, because of the capability to obtain the same information with less SNPs genotyping, thus saving costs and time (Selinski *et al.*, 2011; Hein and Doll, 2012).

In recent studies, the SA phenotype has been associated with a higher risk of developing drug-induced hepatotoxicity than RA or even IA in tuberculosis treatments with INH (Teixeira *et al.*, 2011; Wang *et al.*, 2012; Teixeira *et al.*, 2013). Additionally, the low acetylation also affects the metabolism of acetylhydrazine into non-toxic metabolite, diacetylhydrazine, increasing the rate of accumulation of INH toxic

metabolites. It has been also shown that in SAs the serum concentration of hydrazine was significantly higher than in RAs, possibly due to increased INH concentration (Teixeira *et al.*, 2013).

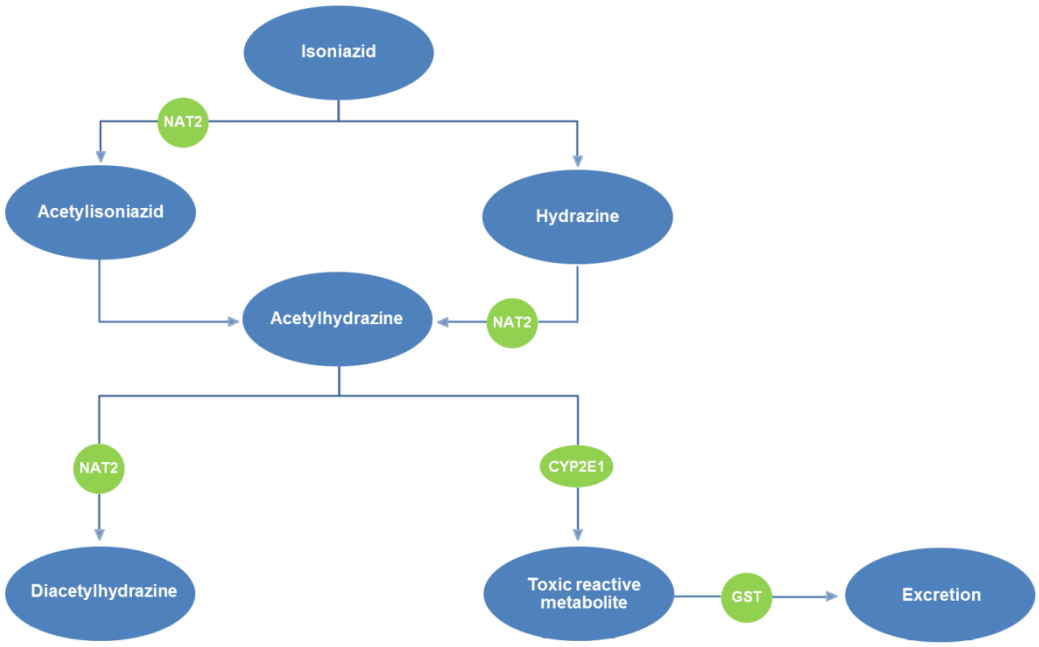


Figure 6. Isoniazid metabolism. This drug is biotransformed into acetylisoniazid by NAT2, hydrolysed into acetylhydrazine and posteriorly oxidised by CYP2E1, forming toxic metabolites. These metabolites are conjugated with glutathione S-transferase (GST) and then excreted. In another metabolic pathway of isoniazid, hydrazine is biotransformed by NAT2 into acetylhydrazine that subsequently is oxidised by CYP2E1. Furthermore, NAT2 also participates in the biotransformation of acetylhydrazine into diacetylhydrazine, a harmless component. Adapted from Teixeira *et al*. (2011), Rios-Santos and Magno (2012) and Teixeira *et al*. (2013).

# 3. Drug targets

Despite the absence of a tight definition for what drug targets are, according to Imming *et al.* (2006) they are molecular structures that pass through a specific interaction with a drug (target-drug interactions) administered to treat a disease (Imming *et al.*, 2006).

Several medicines need to interact with particular targets to exert their pharmacologic effect. When polymorphisms occur in genes that codify these targets, the sensibility to selected drugs may be influenced and consequently it may affect the response of the patient to the treatment. These polymorphisms can have a significant relevance in cases where interindividual variations in drug plasma concentrations are minimal, but major pharmacodynamic differences can be detected (Kupiec *et al.*, 2006; Musshoff *et al.*, 2010; Kudzi *et al.*, 2011).

An example of a drug target polymorphism is the warfarin, which acts by the inhibition of the VKORC1 (Kupiec *et al.*, 2006; Yoshizawa *et al.*, 2009).

## 3.1 Vitamin K epoxide reductase complex subunit 1 – VKORC1

The complex 1 of vitamin k epoxide reductase (VKORC1), is essential in the vitamin k recycling cycle (Markatos *et al.*, 2008; Kwon *et al.*, 2011). The enzyme is responsible for the transformation of vitamin k in its reduced (active) form which is indispensable for the post-translational gamma-carboxylation of vitamin K-dependent clotting factors (Miao *et al.*, 2007).

VKORC1 is the target of several coumarin-based anticoagulant drugs, being warfarin the most frequently prescribed for prevention of thrombotic diseases (Miao *et al.*, 2007; Wang *et al.*, 2008). This anticoagulant drug exerts its pharmacologic effect by inhibiting the enzyme VKORC1, therefore affecting the synthesis of vitamin K-dependent clotting factors (Limdi and Veenstra, 2008; Kwon *et al.*, 2011) (Figure 7).

In the encoding gene, *VKORC1*, located on chromosome 16p12-21q, several genetic variants that alter the sensibility to warfarin have been reported (Montes *et al.*, 2006; Yoshizawa *et al.*, 2009; Scibona *et al.*, 2012).

The c.-1639G>A (rs9923231) polymorphism, the most common variant in Caucasian populations, has been the most widely studied for warfarin and acenocoumarol drugs (Wu, 2007; Markatos *et al.*, 2008).

This polymorphism is associated with a reduction of mRNA expression, with consequent reduction of the VKORC1 activity and warfarin dose requirements (Kwon *et al.*, 2011). Thus, patients with c.-1639AA genotype, require a lower dose of warfarin than those with c.-1639GG genotype (Wu, 2007; Kwon *et al.*, 2011; Smires *et al.*, 2012).

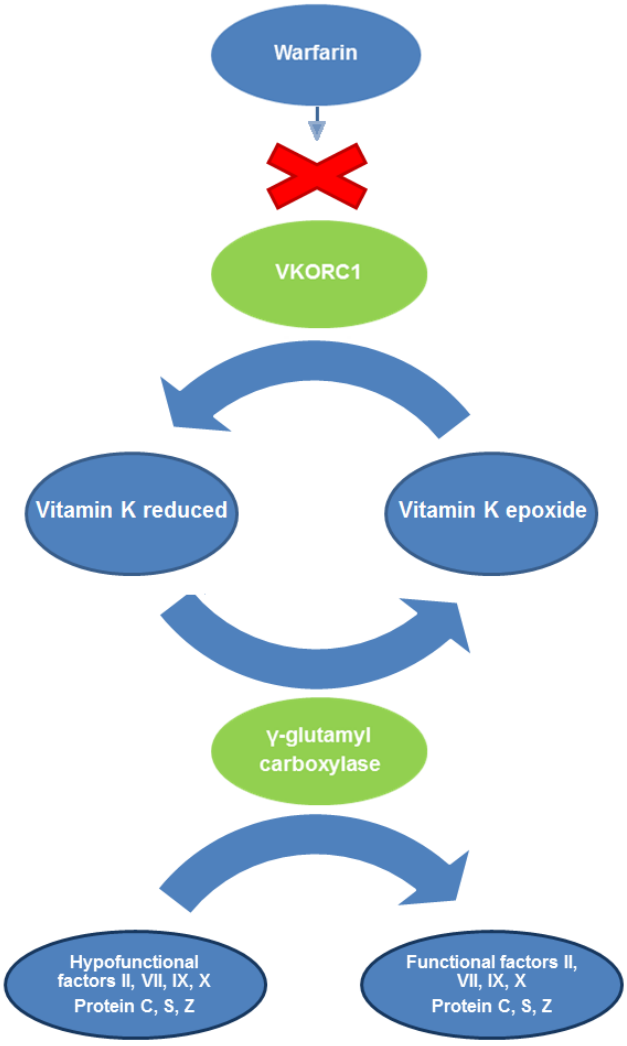


Figure 7. Warfarin effects as vitamin K anticoagulant. This drug interferes with the cyclic inter-conversation of vitamin K, reducing the coagulant activity. Adapted from D'Andrea *et al.* (2008) and Limdi and Veenstra, (2008).

# 4. Worldwide population distribution of pharmacogenetically relevant polymorphisms

If Pharmacogenetics deals with drug metabolism phenotypes which are partly determined by inherited variations in genes encoding for enzymes, transporters and targets involved in the process of drug metabolism, there is no doubt that a main founder pillar of the discipline is human genetic diversity. Along the past decades many studies have revealed that the frequency of alleles influencing drug response exhibit a great variation among different population groups and even within the same population group (Kupiec *et al.*, 2006; Musshoff *et al.*, 2010; Kudzi *et al.*, 2011). For example, the frequency of the CYP2C19 PMs is lower in Western Caucasians than that observed among Orientals (Musshoff *et al.*, 2010). On the other hand, *CYP2C9*2* allele is rarely detected in Asian populations, while *CYP2C9*3* is reported with a frequency at 1.1-6.8% in this population (Kudzi *et al.*, 2011).

Acknowledging this information, Pharmacogenetics has grown hand in hand with extensive population studies aimed at characterising adequately each different population. European populations have been the centre of many of these investigations, leading to obtain a reasonable picture on patterns of pharmacogenetic diversity in Europe. Still, in this regard there are some populations that were poorly studied until now, such as the Gypsy groups.

# 5. Portuguese Gypsies

In the second half of 15th century, coming from the trans-Pyrenees route arrived at Iberia the founding populations that today are known as Gypsies in Portugal. The almost absence of historical records about the ancient people migrations from which would arose the Gypsies, does not allow specifying the origin of these populations, however, evidence coming first from linguistics and more recently from genetic studies both point to India as the region from where departed the proto-Roma populations (Gusmão *et al.*, 2008; Mendizabal *et al.*, 2011).

As Gusmão *et al*. (2008) highlight in their research, while Roma as a whole behave as a transnational genetic isolate, each Roma population presents its specificities, as was observed in Portuguese Gypsies. Roma groups normally have a small population size, with a strong social and cultural identity, where endogamous practices predominate (Gresham *et al.*, 2001). From a genetic viewpoint this leads to the occurrence of high drift effects, what induced the differentiation between the gypsy groups and the host populations, reduced genetic diversity, high linkage disequilibrium and a limited gene flow between Gypsy groups (Gusmão *et al.*, 2008; Valente *et al.*, 2009). Accordingly, no Gypsy group can be genetically representative of others. For these reasons, the Roma have attracted attention from diverse areas of study like, among others, Epidemiology, which would lead to identify a number of Mendelian disorders caused by private founder mutations, and Population Genetics, which is giving important contributions to the reconstruction of the Roma history (Kalaydjieva *et al.*, 2001; Gusmão *et al.*, 2010; Fehér *et al.*, 2011).

In the field of Pharmacogenetics, very few studies were performed up to date focusing on Roma populations, which justifies having elected in this study the Portuguese Gypsies to be analysed from the pharmacogenetic point of view, attempting to understand how much they differ from the host population in which regard genetic polymorphisms influence drug response.

# II. AIMS

Although the Portuguese Gypsies have been demonstrated to be quite differentiated from the Portuguese host population, they remain virtually unstudied in the context of Pharmacogenetics and Forensic Toxicology. Thus, the aim of this study was to characterise the Portuguese Gypsies for single nucleotide polymorphisms (SNP) known to have pharmacogenetic implications, in order to evaluate whether they show any peculiarity in which respect the assessed fraction of diversity that influences drug response.

To achieve this main goal, the following specific aims were established:

1. Selection of SNPs with pharmacogenetic relevance and development of PCR and minisequencing multiplex reactions;
2. Characterisation of a sample of Portuguese Gypsies and another of Portuguese non-Gypsies for the selected SNPs;
3. Comparison of the results obtained in the two population samples;
4. Analysis in a worldwide context of the results obtained in this study.

# III. MATERIAL AND METHODS

## 1. Samples and DNA extraction

For the purpose of the present study, a sample consisting of 116 individuals who self-identified as Portuguese Gypsies was analysed. To perform comparative analyses, a control group of 70 Portuguese non-Gypsy individuals was also studied.

All samples have been extracted for previous works by the Chelex®-100™ (BioRad) method described by Lareu *et al.* (1994).

# 2. Multiplex PCR Design

## 2.1 Selection of Target Polymorphisms

In the beginning of this project, a careful bibliographic analysis was conducted to select genes with pharmacogenetic relevance and potential forensic applications. A total of 5 genes have been selected: *CYP2C9* and *CYP2C19* that encode important enzymes of the phase I of the drug metabolism; *TPMT* and *NAT2,* which codify relevant enzymes for phase II, and *VKORC1* that codify a drug target. For the selection of the variations to be studied, the impact caused to protein activity and reported allelic frequencies in Caucasian populations have been taken into consideration.

From the *CYP2C9* gene were selected *CYP2C9*2* and *CYP2C9*3* allelic variants, that represent about 85% of PM phenotypes and from *CYP2C19* was chosen *CYP2C19*2*, which is relatively common in Caucasian populations and accounts for 75-83% of PM phenotypes (Bozina *et al.*, 2003; Bravo-Villalta *et al.*, 2005). On the other hand, 4 SNPs from the *TPMT* gene were selected, *TPMT*2*, *TPMT*3B*, *TPMT*3C* and *TPMT*8*, being the first 3 responsible for 80-95% of IM and DM phenotypes and 2 SNPs of the *NAT2* gene, c.282C>T and c.341T>C, which were chosen based on their accuracy to predict the NAT2 acetylator phenotypes (Garat *et al.*, 2008; Selinski *et al.*, 2011). At last, *VKORC1* c.-1639G>A was selected due to its significance in drug-target interaction (Smires *et al.*, 2012).

## 2.2 Multiplex PCR amplification

The multiplex PCR has been developed to amplify, at the same time, the various fragments containing the target SNPs.

In a previous study, a multiplex PCR for *TPMT* gene was designed, being in the present work enriched the with new amplification primers for the remaining genes of interest.

The gene sequences were obtained from Ensemble Genome Browser (Ensembl release 68 – July 2012) (http://www.ensembl.org/index.html) and Primer 3 v.0.4.0 Software (http://frodo.wi.mit.edu/) was used to design the amplification primers (Rozen and Skaletsky, 2000), preventing annealing in polymorphic regions. Furthermore, these primers were also designed according to the melting temperature (Tm) of the previous

multiplex PCR, 59ºC, to ensure the action of all primers at the same conditions. Each of the primers obtained by Primer 3 was tested in BLAT (UCSC) (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start), selecting only the human genome, in order to detect possible non-specific annealing, which could reduce the efficiency of the reaction. Then each pair of primers was also tested in UCSC In-Silico PCR (http://genome.ucsc.edu/cgi-bin/hgPcr?command=start) to detect possible non-specific amplifications.

The last step of multiplex development consisted in testing with AutoDimer Software v1.0 and Oligocalc (http://www.basic.northwestern.edu/biotools/OligoCalc.html) if the designed amplification primers could form primer-dimer structures and/or hairpins, which could contribute to reduce the reaction efficiency (Vallone and Butler, 2004). Detailed information about amplification primers, amplified products and the selected SNPs is present in Table 1.

Table 1. Amplification primers and their characteristics.

| Gene | SNP | Allele | Forward primer 5'-3' (bp) | Reverse primer 5'-3' (bp) | Product Size |
|---|---|---|---|---|---|
| *TPMT* | rs1800462 | c.238G>C | TCTTTGAAACCCTATGAACCTGA | TGCGTGCTAAATAGGAACCAT | 377 |
| *TPMT* | rs1800460 | c.460G>A | TGTTGAAGTACCAGCATGCAC | AGCCTTACACCCAGGTCTCT | 365 |
| *TPMT* | rs1142345 | c.719A>G | GAATCCCTGATGTCATTCTTCA | CCTCAAAAACATGTCAGTGTGA | 213 |
| *TPMT* | rs56161402 | c.644G>A | | | |
| *CYP2C9* | rs1799853 | c.430C>T | GGAGGATGGAAAACAGAGACTTACA | AAGGTCAGTGATATGGAGTAGGG | 309 |
| *CYP2C9* | rs1057910 | c.1075A>C | TACCCCTGAATTGCTACAACAAA | GAATTTAATGTCACAGGTCACTGC | 238 |
| *CYP2C19* | rs4244285 | c.681G>A | CAACCAGAGCTTGGCATATTG | TAAAGTCCCGAGGGTTGTTGAT | 223 |
| *NAT2* | rs1041983 | c.282C>T | AGAGGCTATTTTTGATCACATTG | GTGTTTCTTCTTTGGCAGGAGAT | 403 |
| *NAT2* | rs1801280 | c.341T>C | | | |
| *VKORC1* | rs9923231 | c.-1639G>A | GGAGAGGGAAATATCACAGACG | CTCCTGACCTCAAGTGATCCA | 189 |

After the *in silico* evaluation, each primers pair was tested in individual amplification reactions, to confirm if they operate correctly.

Before the multiplex amplification reaction, a mix including all forward and reverse primers was prepared, in which each primer was in a final concentration of 2 µM. The multiplex PCR reactions were performed in a final volume of 10 µL, containing 5 µL of

MyTaq™ HS Mix (Bioline), 1 µL of Primer Mix, 1 µL of DNA and 3 µL of deionized water. The amplification was performed at 95ºC for 15 min, followed by 5 cycles at 94ºC for 30 sec, 63ºC for 90 sec, 72ºC for 1 min, and another 30 cycles at 94ºC for 30 sec, 59ºC for 90 sec, 72ºC for 1 min and a final extension at 72ºC for 10 min.

The amplification reactions were done in 2720 Thermal Cycler (Applied Biosystems) or/and Thermal Cycler (BioRad).

## 2.3 Electrophoresis

Polyacrylamide gel (T9C5) electrophoresis was performed for all samples in order to test the success of the amplification reactions and also identify possible contaminations. The polyacrylamide gel was stained following the Silver staining procedure (Qu *et al.*, 2005).

# 3. SNP genotyping

The genotypic data from the 10 SNPs in study was obtained through the SNaPshot or minisequencing technique described below.

## 3.1 Multiplex Single Base Extension

To perform the minisequencing reactions, 10 single base extension (SBE) primers were designed following the same steps used for amplification primers, but ending immediately adjacent to the target SNPs.

For each SNP, 2 possible extension primers could be designed, forward and reverse, but only one was selected based on the best conditions for the minisequencing reaction according to the Software Primer3. The SBE primers of *CYP2C9* SNPs have a degenerated position (indicated in red at the table) because they anneal with polymorphic sites that could lead to preferential amplification of one of the alleles.

To produce differences in electrophoretic mobility, a non-annealing tail with different sizes was added to the SBE primers. The SBE primers and their characteristics are present in Table 2.

Table 2. Single Base Extension Primers and their characteristics.

| Gene | SNP | Allele | SBE primer sequences 5'-3' | Length | Final length | Detection |
|---|---|---|---|---|---|---|
| *TPMT* | rs1800462 | c.238G>C | AATGTATGATTTTATGCAGGTTT | 23 | 23 | G/C |
| *TPMT* | rs1800460 | c.460G>A | TGACATGATTTGGGATAGAGGA | 22 | 31 | G/A |
| *TPMT* | rs1142345 | c.719A>G | TGTCTCATTTACTTTTCTGTAAGTAGA | 27 | 36 | T/C |
| *TPMT* | rs56161402 | c.644G>A | TTTCAGGTAAAATATGCAATATAC | 24 | 48 | G/A |
| *CYP2C9* | rs1799853 | c.430C>T | GGAAGAGGAGCATTGAGGA**Y** | 20 | 53 | C/T |
| *CYP2C9* | rs1057910 | c.1075A>C | GTGCACGAGGTCCAGAG**R**TAC | 21 | 58 | A/C |
| *CYP2C19* | rs4244285 | c.681G>A | GTTTTTAAGTAATTTGTTATGGGTTCC | 27 | 68 | C/T |
| *NAT2* | rs1041983 | c.282C>T | GTGCTGTATTTGTTAACTGGAGGGAT | 26 | 63 | G/A |
| *NAT2* | rs1801280 | c.341T>C | CCTTCTCCTGCAGGTGACCA | 20 | 42 | T/C |
| *VKORC1* | rs9923231 | c.-1639G>A | GAAAAACAACCATTGGCC | 18 | 18 | G/A |

## 3.2 Minisequencing protocol

Before the minisequencing reactions, an initial purification with ExoSAP-IT® was performed in order to remove the excess of dNTPs and primers present in the PCR product. Each purification reaction contained 1 µL of amplified product and 0.8 µL of ExoSAP-IT® (USB® Products) and the protocol used was 37ºC for 15 min and 85ºC for another 15 min.

A mix including all SBE primers was prepared, containing 0.4 µL of the *VKORC1* SBE primer and 0.2 µL of the remaining SBE primers, all with an initial concentration of 20 µM.

The SNaPshot™ reactions were performed in a final volume of 5 µL using 1.8 µL of purified PCR product, 2.2 µL of SBE primers mix and 1 µL of SNaPshot™ Multiplex Kit (Applied Biosystems). The reaction conditions were 25 cycles at 96ºC for 10 sec, 50ºC for 5 sec and 60ºC for 30 sec.

The final products of the SNaPshot reactions were treated with 1 µL of SAP enzyme (USB®) at 37ºC for 60 min and 85ºC for 15 min, to eliminate unincorporated ddNTPs.

These reactions were carried out in a 2720 Thermal Cycler (Applied Biosystems) or/and Thermal Cycler (BioRad).

Finally, 3 µL of the minisequencing-purified products were mixed with 12 µL of Hi-Di™ Formamide (Applied Biosystems) and GeneScan™-120 LIZ™ size standard. Capillary electrophoresis was then performed in an ABI 3130 Genetic Analyser (Applied Biosystems).

# 4. Sanger Sequencing

The Sanger sequencing technique was performed to validate the results obtained in multiplex PCR and SNaPshot reactions.

An initial purification with ExoSAP-IT® was performed, such as in minisequencing procedure, to eliminate the elements that were not consumed in the amplification reaction. Each purification reaction contained 1.5 µL of amplified product and 0.5 µL of ExoSAP-IT®. The reaction conditions were described in section 3.2.

The sequencing reactions were prepared in a total volume of 5 µL, containing 2 µL of purified PCR product, 1 µL of the amplification primer (forward or reverse) at 2 µM, 1 µL of BigDye sequencing kit and 1 µL of sequencing buffer. The protocol used was 96ºC for 4 min, followed by 35 cycles at 96ºC for 10 sec, 55ºC for 5 sec and 60ºC for 4 min, and a final step at 60ºC for 10 min.

The reactions were executed in 2720 Thermal Cycler (Applied Biosystems) or/and Thermal Cycler (BioRad).

Then, the sequencing products were purified in Sephadex® columns (750 µL) through centrifugation for 4 minutes at 4400 rpm.

The final products were resuspended in 12 µL of Hi-Di™ Formamide and capillary electrophoresis was performed in an ABI 3130 Genetic Analyser (Applied Biosystems).

# 5. Data analysis

The results of SNaPshot were analysed with GeneMapper® v4.0 and Peak Scanner™ v1.0 Softwares and the sequencing data was analysed with Sequencing Analysis v5.2 and Geneious Pro 5.5.8.

The Arlequin Software v3.5 was used to determine the allelic frequencies, to test departures from Hardy-Weinberg Equilibrium and to perform the Fisher's exact test (Excoffier and Lischer, 2010). The significant level assumed for these tests was adjusted applying the Bonferroni correction.

The *STATISTICA* 10 software was also used to execute the Principal Components & Classification Analysis (Version 10, 2010).

# IV. RESULTS AND DISCUSSION

## 1. Locus by locus approach

The PCR and minisequencing multiplex reactions were developed with success. In figure 8 is observed an example of the pattern bands of the PCR multiplex reaction in polyacrylamide gel and in figure 9 is observed an example of an electropherogram.


Figure 8. Polyacrylamide gel. Band patterns observed after the electrophoresis and respective fragments length and SNPs.

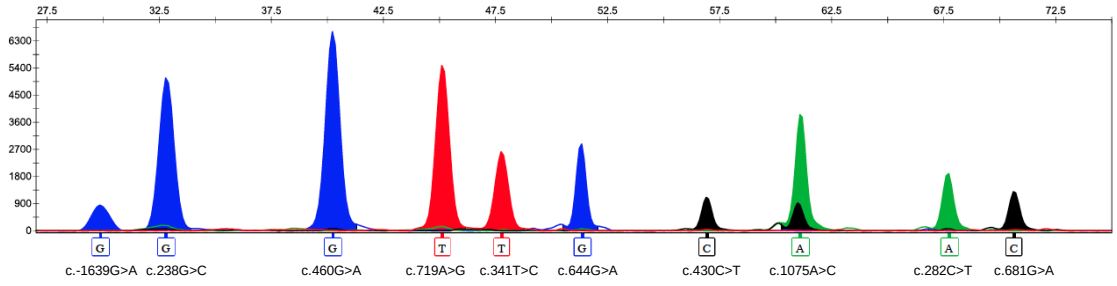


Figure 9. Electropherogram obtained after multiplex minisequencing analysis. For the polymorphisms c.719A>G, c.282C>T and c.681G>A were designed reverse extension primers, leading to the detection of the complementary base.

The genotypic distributions observed for each of the scrutinized variations were tested for Hardy-Weinberg Equilibrium with the Arlequin software. No significant deviations were found in the Gypsies or in the non-Gypsies samples (Supplementary Table 1).

## 1.1 Cytochrome P450 2C9 - CYP2C9

As previously referred, the CYP2C9 is an important metabolic enzyme responsible for the metabolism of diverse drugs, some of them with a narrow therapeutic index, such as Warfarin (Xie *et al.*, 2002).

The two selected SNPs*, CYP2C9*2* and *CYP2C9*3*, are responsible for a reduction of the enzymatic capacity of CYP2C9 and the individuals that carry one or both of these variants can more frequently develop ADRs than the individuals that carry only wild-type alleles (*CYP2C9*1*) (Carlquist *et al.*, 2006; Yousef *et al.*, 2012).

In this study, the allelic frequencies of *CYP2C9*2* and *CYP2C9*3* in the Portuguese non-Gypsy and Gypsy populations were estimated and are presented in Table 3.

Table 3. Allele frequencies of *CYP2C9* in the Portuguese non-Gypsy and Gypsy populations.

| **CYP2C9** | **Portuguese non-Gypsies (n)** | **Portuguese Gypsies (n)** |
|---|---|---|
| *CYP2C9*2* | 0.200±0.035 (70) | 0.108±0.020* (116) |
| *CYP2C9*3* | 0.057±0.020 (70) | 0.079±0.018 (114) |

*Significant value ($p \leq 0.05$)
n, sample size

To assess whether differences between frequencies of *CYP2C9*2* and *CYP2C9*3* in the two populations were statistically significant, Fisher exact tests of population differentiation were performed, revealing a significant difference for *CYP2C9*2* (p=0.02070±0.0018), whereas, no significant difference was found for *CYP2C9*3* (p= 0.53454±0.0027).

Next, comparisons were extended to diverse worldwide populations, for which data had been previously published and collected from the bibliography. Among them was included another sample from Portugal that had been previously characterised by

Oliveira *et al*. (2007a). The Fisher exact test was also applied in these comparisons to evaluate if the differences observed between allelic frequencies had statistical significance, using the significance level of inferred after the application of the Bonferroni correction for multiple tests. The p-values obtained in the comparisons are present in Table 4 for *CYP2C9*2* and in Table 5 for *CYP2C9*3*.

In general, the allele frequency of *CYP2C9*2* in the Portuguese non-Gypsies had significant differences comparatively to values found in most of the African and Asian populations. In contrast, when it was compared with frequencies reported for European populations the majority of the differences had no statistical significance, as well as did not differences with the available population samples from America, among which were not included representatives of Native-Americans.

Similar results were observed when the allele frequency of *CYP2C9*2* in the Portuguese Gypsies was compared with worldwide data.

For *CYP2C9*3*, the Portuguese non-Gypsy sample only showed significant differences with Benin, Ghana and the Roma from Hungary, whereas the Portuguese Gypsies significantly differed from Benin, Ghana, USA African-Americans and Japan.

As shown in Table 4 and 5, Spikey *et al*. (2009b) had provided data for the Roma from Hungary regarding to the two *CYP2C9* alleles here examined. Interestingly, when compared to the Portuguese Gypsies, the Hungarian Roma did not showed differences at *CYP2C9*2* and *CYP2C9*3*, whereas comparatively to the Portuguese non-Gypsies they differed significantly at *CYP2C9*3* allele.

A special attention was also given to the comparisons between the Portuguese Gypsies and Indian populations, since linguistic and genetic evidence indicate that the Northern region of the Indian sub-continent, was likely the original homeland of the migrants from whom the current Roma arose (Pereira *et al.*, 2012). In the studies of Nahar *et al*. (2013) and Jose *et al*. (2005) were also estimated *CYP2C9*2* and *CYP2C9*3* frequencies in North and South India, respectively. When the allelic frequencies of the Portuguese Gypsies were compared with values for North India, differences were not found to be statistically significant for any of the two alleles. However, the frequency of *CYP2C9*2* in the Portuguese Gypsies significantly differed from that reported for South India.

Table 4. Comparison of allele frequencies of *CYP2C9*2* reported from different ethnic populations.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Africa | | | | | | |
| | Benin | 111 | 0 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Allabi *et al.*, 2003) |
| | Egypt | 247 | 0.120 | 0.01564±0.0017* | 0.71072±0.0048 | (Hamdy *et al.*, 2002) |
| | Ethiopia | 150 | 0.043 | 0.00000±0.0000*† | 0.00706±0.0007* | (Scordo *et al.*, 2001) |
| | Ghana | 204 | 0 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Kudzi *et al.*, 2009) |
| | South Africa | 993 | 0 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Dandara *et al.*, 2011) |
| America | | | | | | |
| | African-American | 100 | 0.010 | 0.00000±0.0000*† | 0.00003±0.0000*† | (Hamdy *et al.*, 2002) |
| | Argentina | 101 | 0.262 | 0.19927±0.0088 | 0.00002±0.0000*† | (Scibona *et al.*, 2012) |
| | Ashkenazi Jewish | 510 | 0.127 | 0.02591±0.0014* | 0.43004±0.0096 | (Scott *et al.*, 2008) |
| | Bolivia | 778 | 0.048 | 0.00000±0.0000*† | 0.00032±0.0001*† | (Bravo-Villalta *et al.*, 2005) |
| | Sephardi Jewish | 80 | 0.194 | 1.00000±0.0000 | 0.01714±0.0017* | (Scott *et al.*, 2008) |
| | USA | 140 | 0.132 | 0.08654±0.0050 | 0.41993±0.0088 | (Dickmann *et al.*, 2001) |
| Asia | | | | | | |
| | China | 394 | 0.001 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Yang *et al.*, 2003) |
| | Iran | 150 | 0.253 | 0.23076±0.0025 | 0.00001±0.0000*† | (Azarpira *et al.*, 2010) |
| | Japan | 341 | 0 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Yoshizawa *et al.*, 2009) |
| | Jordania | 263 | 0.135 | 0.05817±0.0028 | 0.34428±0.0072 | (Yousef *et al.*, 2012) |
| | Korea | 358 | 0 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Bae *et al.*, 2005) |
| | Lebanon | 146 | 0.113 | 0.01564±0.0017* | 0.89184±0.0021 | (Saad and Langaee, 2011) |
| | Malaysia | 85 | 0.006 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Gan *et al.*, 2011) |
| | North India | 209 | 0.050 | 0.00000±0.0000*† | 0.00924±0.0012* | (Nahar *et al.*, 2013) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | Saudi Arabia | 131 | 0.133 | 0.08004±0.0026 | 0.41258±0.0081 | (Alzahrani *et al.*, 2013) |
| | South India | 346 | 0.040 | 0.00000±0.0000 *† | 0.00038±0.0002 *† | (Jose *et al.*, 2005) |
| | Turkey | 85 | 0.100 | 0.01589±0.0014* | 0.87448±0.0017 | (Babaoglu *et al.*, 2004) |
| Europe | | | | | | |
| | Belgium | 121 | 0.100 | 0.00815±0.0009* | 0.76622±0.0048 | (Allabi *et al.*, 2003) |
| | Croatia | 200 | 0.165 | 0.34634±0.0068 | 0.05097±0.0034 | (Bozina *et al.*, 2003) |
| | Denmark | 276 | 0.121 | 0.01623±0.0016* | 0.62487±0.0068 | (Pedersen *et al.*, 2010) |
| | Faroe Islands | 311 | 0.088 | 0.00044±0.0002 *† | 0.43034±0.0099 | (Pedersen *et al.*, 2010) |
| | France | 151 | 0.150 | 0.21672±0.0088 | 0.20514±0.0078 | (Yang *et al.*, 2003) |
| | Germany | 118 | 0.140 | 0.14210±0.0045 | 0.33511±0.0041 | (Burian *et al.*, 2002) |
| | Greece | 283 | 0.129 | 0.04065±0.0039 * | 0.47106±0.0053 | (Arvanitidis *et al.*, 2007) |
| | Hungary | 535 | 0.125 | 0.01698±0.0021* | 0.50442±0.0079 | (Sipeky *et al.*, 2009b) |
| | Hungary (Gypsy) | 465 | 0.118 | 0.01021±0.0016* | 0.72840±0.0037 | (Sipeky *et al.*, 2009b) |
| | Italy | 360 | 0.125 | 0.02355±0.0032* | 0.55310±0.0081 | (Scordo *et al.*, 2004) |
| | Norway | 328 | 0.099 | 0.00125±0.0004 * | 0.71109±0.0070 | (Pedersen *et al.*, 2010) |
| | Portugal | 125 | 0.132 | 0.08395±0.0052 | 0.47511±0.0039 | (Oliveira *et al.*, 2007a) |
| | Romania | 332 | 0.113 | 0.00935±0.0016* | 0.90333±0.0024 | (Buzoianu *et al.*, 2012) |
| | Russia | 290 | 0.105 | 0.00317±0.0007* | 0.90394±0.0015 | (Gaikovitch *et al.*, 2003) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | Slovenia | 127 | 0.122 | 0.05245±0.0017 | 0.67028±0.0057 | (Herman *et al.*, 2003) |
| | Spain | 200 | 0.120 | 0.02276±0.0027* | 0.69737±0.0053 | (Mas *et al.*, 2005) |
| | Sweden | 1490 | 0.113 | 0.00508±0.0010* | 0.91801±0.0018 | (Wadelius *et al.*, 2009) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.001
n, sample size
q, frequencies of the minor allele
Ref., references

Table 5. Comparison of allele frequencies of *CYP2C9*3* reported from different ethnic populations.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Africa | | | | | | |
| | Benin | 111 | 0 | 0.00061±0.0002*† | 0.00000±0.0000*† | (Allabi *et al.*, 2003) |
| | Egypt | 247 | 0.060 | 1.00000±0.0000 | 0.40683±0.0071 | (Hamdy *et al.*, 2002) |
| | Ethiopia | 150 | 0.023 | 0.09045±0.0024 | 0.00312±0.0007* | (Scordo *et al.*, 2001) |
| | Ghana | 204 | 0 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Kudzi *et al.*, 2009) |
| America | | | | | | |
| | African-American | 100 | 0.005 | 0.00375±0.0005* | 0.00009±0.0000*† | (Hamdy *et al.*, 2002) |
| | Argentina | 101 | 0.030 | 0.26901±0.0039 | 0.03329±0.0016* | (Scibona *et al.*, 2012) |
| | Ashkenazi Jewish | 510 | 0.081 | 0.40291±0.0064 | 1.00000±0.0000 | (Scott *et al.*, 2008) |
| | Bolivia | 778 | 0.030 | 0.12696±0.0020 | 0.00114±0.0005* | (Bravo-Villalta *et al.*, 2005) |
| | Sephardi Jewish | 80 | 0.144 | 0.02086±0.0019* | 0.04227±0.0034* | (Scott *et al.*, 2008) |
| | USA | 140 | 0.043 | 0.62699±0.0046 | 0.09765±0.0046 | (Dickmann *et al.*, 2001) |
| Asia | | | | | | |
| | China | 394 | 0.036 | 0.23142±0.0033 | 0.00927±0.0006* | (Yang *et al.*, 2003) |
| | Iran | 150 | 0.098 | 0.19955±0.0053 | 0.53338±0.0048 | (Azarpira *et al.*, 2010) |
| | Japan | 341 | 0.021 | 0.03750±0.0014* | 0.00022±0.0001*† | (Yoshizawa *et al.*, 2009) |
| | Jordania | 263 | 0.068 | 0.71311±0.0032 | 0.64234±0.0040 | (Yousef *et al.*, 2012) |
| | Korea | 358 | 0.060 | 1.00000±0.0000 | 0.35730±0.0080 | (Bae *et al.*, 2005) |
| | Lebanon | 146 | 0.096 | 0.18839±0.0065 | 0.53423±0.0079 | (Saad and Langaee, 2011) |
| | Malaysia | 85 | 0.041 | 0.60368±0.0026 | 0.14645±0.0026 | (Gan *et al.*, 2011) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | North India | 209 | 0.110 | 0.06837±0.0043 | 0.21124±0.0065 | (Nahar *et al.*, 2013) |
| | Saudi Arabia | 131 | 0.023 | 0.08866±0.0020 | 0.00572±0.0011* | (Alzahrani *et al.*, 2013) |
| | South India | 346 | 0.080 | 0.48582±0.0027 | 1.00000±0.0000 | (Jose *et al.*, 2005) |
| | Turkey | 85 | 0.088 | 0.38896±0.0052 | 0.85456±0.0026 | (Babaoglu *et al.*, 2004) |
| Europe | | | | | | |
| | Belgium | 121 | 0.074 | 0.67586±0.0030 | 0.86319±0.0020 | (Allabi *et al.*, 2003) |
| | Croatia | 200 | 0.095 | 0.21493±0.0042 | 0.56737±0.0041 | (Bozina *et al.*, 2003) |
| | Denmark | 276 | 0.053 | 0.83493±0.0020 | 0.17889±0.0047 | (Pedersen *et al.*, 2010) |
| | Faroe Islands | 311 | 0.053 | 0.83322±0.0019 | 0.18491±0.0034 | (Pedersen *et al.*, 2010) |
| | France | 151 | 0.080 | 0.43940±0.0065 | 1.00000±0.0000 | (Yang *et al.*, 2003) |
| | Germany | 118 | 0.050 | 0.81451±0.0010 | 0.25619±0.0056 | (Burian *et al.*, 2002) |
| | Greece | 283 | 0,081 | 0.37836±0.0071 | 1.00000±0.0000 | (Arvanitidis *et al.*, 2007) |
| | Hungary | 535 | 0.088 | 0.25441±0.0062 | 0.79412±0.0035 | (Sipeky *et al.*, 2009b) |
| | Hungary (Gypsy) | 465 | 0.155 | 0.00085±0.0002*† | 0.00158±0.0004* | (Sipeky *et al.*, 2009b) |
| | Italy | 360 | 0.097 | 0.14524±0.0041 | 0.50615±0.0057 | (Scordo *et al.*, 2004) |
| | Norway | 328 | 0.065 | 0.84796±0.0017 | 0.53567±0.0056 | (Pedersen *et al.*, 2010) |
| | Portugal | 129 | 0.080 | 0.42545±0.0069 | 1.00000±0.0000 | (Oliveira *et al.*, 2007a) |
| | Romania | 332 | 0.093 | 0.18419±0.0072 | 0.60297±0.0074 | (Buzoianu *et al.*, 2012) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | Russia | 290 | 0.067 | 0.84642±0.0014 | 0.54268±0.0077 | (Gaikovitch *et al.*, 2003) |
| | Slovenia | 127 | 0.063 | 1.00000±0.0000 | 0.59735±0.0025 | (Herman *et al.*, 2003) |
| | Spain | 200 | 0.062 | 1.00000±0.0000 | 0.51289±0.0071 | (Mas *et al.*, 2005) |
| | Sweden | 1490 | 0.071 | 0.61582±0.0050 | 0.70366±0.0068 | (Wadelius *et al.*, 2009) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.001
n, sample size
q, frequencies of the minor allele
Ref., references

Is worthy of note the pattern of frequencies across European, Portuguese Gypsy and Indian populations, observed for *CYP2C9*2*. In the Portuguese Gypsies the frequency is intermediate between the European and Indian values (Figure 10). This finding may be a signal of the long lasting process of admixture that has characterized the history of all Roma groups. A few centuries after having left India, the Roma groups arrived in Europe, where they spread into the entire continent through multiple dispersion routes, during which strong interactions were established with the host populations, as well did when they become established in specific European regions (Kalaydjieva *et al.*, 2005).

Figure 10. *CYP2C9*2* allele frequency distribution of the European, Indian and Gypsy populations. The frequency of the Portuguese Gypsy population is represented in blue, the frequency of the Portuguese non-Gyspy population is represented in red and the frequency of the Hungarian Roma population is represented in violet.

## 1.2 Cytochrome P450 2C19 – CYP2C19

CYP2C19 is another CYP enzyme that plays a significant role in the activation and detoxification of several classes of drugs such as clopidogrel and omeprazole (Zanger *et al.*, 2008).

The *CYP2C19*2* is associated with the PM phenotype, which is characterised by an absence of enzymatic activity. Interestingly, this CYP2C19 PM phenotype seems to have a role in *H. pylori* eradication therapy, because its administration in conjunction with the antibiotics produces, in PM patients compared to other phenotypes, more accentuated improvements on the symptoms and on effectiveness of the antibiotics (Zanger *et al.*, 2008).

The allele frequencies of *CYP2C19*2* estimated for the Portuguese non-Gypsy and Gypsy samples are shown in Table 6. No differences between them were detected through the Fisher exact test (p=0.26675±0.0083).

Table 6. Allele frequencies of *CYP2C19* in the Portuguese non-Gypsy and Gypsy populations.

| CYP2C19 | Portuguese non-Gypsies (n) | Portuguese Gypsies (n) |
|---|---|---|
| *CYP2C19*2* | 0.164±0.031 (70) | 0.211±0.027 (116) |

n, sample size

Comparative analyses were again performed between data here obtained and allelic frequencies recruited from the literature, with differences being assessed as previously referred, according to the correction for multiple tests (Table 7). The frequency of *CYP2C19*2* in the Portuguese non-Gypsy reference sample was only statistically different from values in Ghana, China, South India and Thailand. On the other hand, the frequency of the same allele in the Portuguese Gypsies revealed significant differences comparatively to frequencies in some African populations, but in contrast, it did not differed from the registered in the majority of Asian and European populations.

In the before mentioned study of Oliveira *et al*. (2007a), where *CYP2C19*2* was also screened in a Portuguese sample, the frequency estimate obtained was not statistically different from the values here obtained either in Gypsies or non-Gypsies.

Data available for the Roma group from Hungary (Sipeky *et al.*, 2013), did not revealed differences comparatively to the *CYP2C19*2* frequency in the Portuguese Gypsies, but is of note that both Roma groups present the highest frequencies of *CYP2C19*2* in populations from Europe.

In North and South India *CYP2C19*2* is very well represented (Lamba *et al*. (2000) and Jose *et al*. (2005), respectively), reaching frequencies higher than usually found in European populations. Comparing the allele frequency in the Portuguese Gypsies with that in North India, no statistical difference was found, but on the contrary, when it was compared with the frequency in South India, the difference reached statistical significance.

Table 7. Comparison of allele frequencies of *CYP2C19*2* reported from different ethnic populations.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Africa | | | | | | |
| | Benin | 111 | 0.130 | 0.43008±0.0057 | 0.02567±0.0023* | (Allabi *et al.*, 2003) |
| | Egypt | 247 | 0.110 | 0.10530±0.0040 | 0.00071±0.0008*† | (Hamdy *et al.*, 2002) |
| | Ghana | 204 | 0.060 | 0.00032±0.0001*† | 0.00000±0.0000*† | (Kudzi *et al.*, 2009) |
| | South Africa | 993 | 0.160 | 0.90646±0.0013 | 0.05003±0.0015 | (Dandara *et al.*, 2011) |
| America | | | | | | |
| | African-American | 108 | 0.250 | 0.06503±0.0039 | 0.35077±0.0052 | (Hamdy *et al.*, 2002) |
| | Bolivia | 778 | 0.078 | 0.00132±0.0005* | 0.00000±0.0000*† | (Bravo-Villalta *et al.*, 2005) |
| | Colombia (Mestizos) | 189 | 0.087 | 0.01565±0.0007* | 0.00000±0.0000*† | (Isaza *et al.*, 2007) |
| | Brazil (African descent) | 99 | 0.170 | 0.40362±0.0044 | 0.90194±0.0025 | (Santos *et al.*, 2011) |
| | Brazil (Amerindian) | 183 | 0.104 | 0.07652±0.0037 | 0.00022±0.0001*† | (Santos *et al.*, 2011) |
| | Brazil (Caucasian descent) | 615 | 0.170 | 1.00000±0.0000 | 0.13036±0.0072 | (Santos *et al.*, 2011) |
| | Brazil (Malutto) | 315 | 0.165 | 1.00000±0.0000 | 0.13616±0.0078 | (Santos *et al.*, 2011) |
| | USA | 105 | 0.129 | 0.34890±0.0070 | 0.02489±0.0022* | (Hamdy *et al.*, 2002) |
| Asia | | | | | | |
| | China | 121 | 0.455 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Xiao *et al.*, 1997) |
| | Iran | 150 | 0.130 | 0.36399±0.0071 | 0.01235±0.0015* | (Azarpira *et al.*, 2010) |
| | Japan | 265 | 0.279 | 0.00651±0.0015* | 0.04567±0.0014* | (Sugimoto *et al.*, 2008) |
| | Jordania | 158 | 0.123 | 0.23683±0.0048 | 0.00564±0.0007* | (Yousef *et al.*, 2012) |
| | Korea | 271 | 0.284 | 0.00331±0.0005* | 0.03807±0.0042* | (Kim *et al.*, 2010) |
| | Lebanon | 161 | 0.134 | 0.39725±0.0075 | 0.01813±0.0021* | (Djaffar Jureidini *et al.*, 2011) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | Malaysia | 142 | 0.280 | 0.00707±0.0008* | 0.06476±0.0045 | (Yang *et al.*, 2004) |
| | North India | 121 | 0.297 | 0.00365±0.0007* | 0.03382±0.0025* | (Sipeky *et al.*, 2013) |
| | South India | 341 | 0.350 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Jose *et al.*, 2005) |
| | Thailand | 121 | 0.351 | 0.00010±0.0001*† | 0.00075±0.0006*† | (Yamada *et al.*, 2001) |
| | Turkey | 404 | 0.120 | 0.16792±0.0067 | 0.00035±0.0002*† | (Sipeky *et al.*, 2013) |
| | Vietnam | 90 | 0.236 | 0.15195±0.0063 | 0.64068±0.0063 | (Yamada *et al.*, 2001) |
| Europe | | | | | | |
| | Belgium | 121 | 0.091 | 0.04630±0.0017* | 0.00069±0.0005*† | (Allabi *et al.*, 2003) |
| | Croatia | 200 | 0.150 | 0.69008±0.0035 | 0.06647±0.0043 | (Bozina *et al.*, 2003) |
| | Denmark | 64 | 0.180 | 0.75762±0.0037 | 0.49594±0.0061 | (Bathum *et al.*, 1998) |
| | Faroe Islands | 311 | 0.187 | 0.62184±0.0086 | 0.43595±0.0163 | (Pedersen *et al.*, 2010) |
| | Germany | 140 | 0.150 | 0.77998±0.0042 | 0.07895±0.0042 | (Hamdy *et al.*, 2002) |
| | Greece | 283 | 0.137 | 0.42378±0.0065 | 0.01434±0.0016* | (Arvanitidis *et al.*, 2007) |
| | Hungary | 370 | 0.126 | 0.22115±0.0085 | 0.00113±0.0005* | (Sipeky *et al.*, 2013) |
| | Hungary (Gypsy) | 500 | 0.205 | 0.31091±0.0060 | 0.86004±0.0041 | (Sipeky *et al.*, 2013) |
| | Italy | 360 | 0.111 | 0.08398±0.0036 | 0.00009±0.0001*† | (Scordo *et al.*, 2004) |
| | Norway | 328 | 0.152 | 0.69624±0.0056 | 0.03961±0.0040* | (Pedersen *et al.*, 2010) |
| | Portugal | 126 | 0.140 | 0.54840±0.0057 | 0.04465±0.0031* | (Oliveira *et al.*, 2007a) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | Romania | 200 | 0.137 | 0.48033±0.0074 | 0.01769±0.0017* | (Buzoianu *et al.*, 2010) |
| | Russia | 290 | 0.114 | 0.12036±0.0072 | 0.00031±0.0002*† | (Gaikovitch *et al.*, 2003) |
| | Slovenia | 127 | 0.159 | 0.88272±0.0024 | 0.13727±0.0067 | (Herman *et al.*, 2003) |
| | Spain | 177 | 0.130 | 0.31563±0.0046 | 0.01352±0.0027* | (Arias *et al.*, 2005) |
| | Sweden | 185 | 0.160 | 0.89279±0.0023 | 0.13596±0.0082 | (Ramsjö *et al.*, 2010) |
| Oceania | | | | | | |
| | Australia | 99 | 0.145 | 0.64714±0.0053 | 0.10258±0.0068 | (Hoskins *et al.*, 1998) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.001
n, sample size
q, frequencies of the minor allele
Ref., references

So then, the distribution of the frequencies of *CYP2C19*2* in European, Portuguese Gypsies and Indian populations, follows the trend observed for *CYP2C9*2*, as is illustrated in the map of the Figure 11, which makes quite clear that both the Roma from Hungary and Portugal present *CYP2C19*2* frequencies that are in the transition between those typical of non-Gypsy Europeans and those typical of Indians.

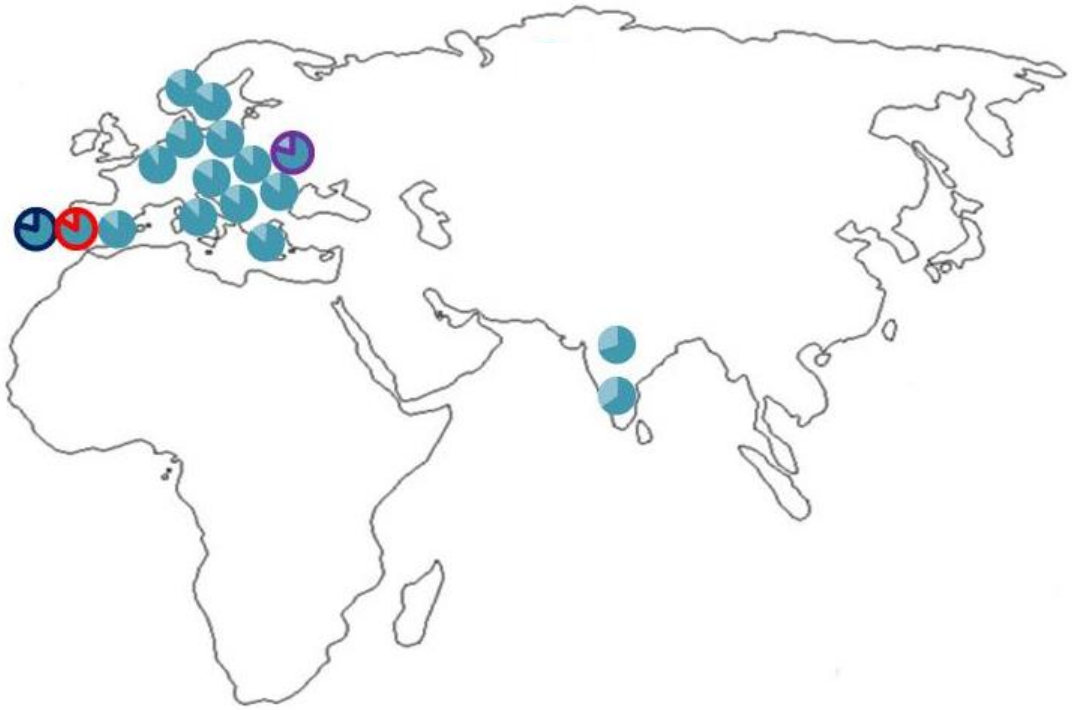

Figure 11. *CYP2C19*2* allele frequency distribution of the European, Indian and Gypsy populations. The frequency of the Portuguese Gypsy population is represented in blue, the frequency of the Portuguese non-Gypsy population is represented in red and the frequency of the Hungarian Roma population is represented in violet.

## 1.3 Thiopurine S-methyltransferase – TPMT

Thiopurine S-methyltransferase (TPMT), a cytoplasmic enzyme that catalyses the methylation of diverse thiopurine drugs, is often considered one of the best models in Pharmacogenetics due to its well established relevance in thiopurine chemotherapy (Kubota and Chiba, 2001; Krynetski and Evans, 2003).

The studied SNPs in the gene encoding for this enzyme define alleles *TPMT*2*, *TPMT*3A*, *TPMT*3C* and *TPMT*8*, which together are responsible for the majority of deficient methylator phenotypes in most populations, even though presenting considerable distribution heterogeneity across major human populations (Oliveira *et al.*, 2007b; Garat *et al.*, 2008; Roberts and Barclay, 2012). For instance, while *TPMT*3A* prevail in Caucasians, *TPMT*3C* is the predominant allele in Asian and African populations and *TPMT*8* has been quite restrictively detected in African populations or populations of African descent (Wang *et al.*, 2010). Accordingly, *TPMT*3C* was not detected in this study. Also *TPMT*2* was absent from both samples examined, which indeed was also not surprising given that *TPMT*2* has been demonstrated to be one of

the less frequent deficient alleles in most populations where it was identified (McLeod *et al.*, 2000).

Out of the referred 4 *TPMT* alleles, only *TPMT*8* and *TPMT*3A* were detected in this study, and the estimated frequencies in Portuguese non-Gypsies and Gypsies are summarised in Table 8.

Table 8. Allele frequencies of *TPMT* in the Portuguese non-Gypsy and Gypsy populations.

| TPMT | Portuguese non-Gypsies (n) | Portuguese Gypsies (n) |
|---|---|---|
| *TPMT*3A* | 0.029±0.014 (70) | 0.004±0.004 (116) |
| *TPMT*8* | 0.000±0.000 (70) | 0.004±0.004 (116) |

n, sample size

*TPMT*8* was only found in a heterozygous Gypsy individual. On the other hand, *TPMT*3A* was detected in both samples, at the low frequency of 2.9% in the Portuguese non-Gypsies and at the very low frequency of 0.4% in the Portuguese Gypsies. Neither for *TPMT*3A* nor for *TPMT*8* differences in frequencies between both samples were statistically significant (p=0.07144±0.0022 for *TPMT*3A* and p= 1.00000±0.0000 for *TPMT*8*).

In Tables 9 and 10 are presented the results of the comparative tests performed between data here obtained for *TPMT*3A* and *TPMT*8*, respectively, and previously published frequencies for other populations.

For *TPMT*3A*, the frequency of 2.9% in the Portuguese non-Gypsies lies in the range of values commonly found in other European populations, where the allele presents its peak of frequency at 8.6% in a sample from Germany. In Africa, *TPMT*3A* has only been sporadically detected and in most Asian populations it occurs at very low frequency. In line with this overall pattern, the frequency of *TPMT*3A* in the Portuguese non-Gypsies was found to be statically different from the typical values in African and many Asian populations assuming the conventional p≤0.05 as level of significance, although when corrected for multiple tests only differences with Korea and South India maintained statistical significance. As for *TPMT*3A* in the Portuguese Gypsies, differences in frequencies were only significant with Bolivian and German considering the correction for multiple tests.

For *TPMT*8* allele, no differences with statistical significance were found, which otherwise is understandable given that this rare deficient allele only reaches frequencies around 2% in African populations, and so due to relative rarity it is difficult to capture statistical differences.

Comparing our *TPMT* data for the Portuguese reference sample with that previously published by Alves *et al.* (2000) and Alves *et al*. (2001) also for Portugal, a strong agreement is observed, despite *Alves et al*. had additionally detected *TPMT*2* at the very low frequency of 1.5% in the study of 2000 and 1% in the study of 2001.

Finally, focusing in the samples from the North and South India studied by Kapoor *et al*. (2009) and Umamaheswaran *et al*. (2012), respectively, *TPMT*3A* was not detected in South India, whereas in the North its presence is almost residual, 0.04%, which coincidentally equals the frequency registered in the Portuguese Gypsies.

Table 9. Comparison of allele frequencies of *TPMT*3A* reported from different ethnic populations.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Africa | | | | | | |
| | African-American | 248 | 0.008 | 0.07907±0.0024 | 1.00000±0.0000 | (Hon *et al.*, 1999) |
| | Cabinda | 103 | 0 | 0.02736±0.0006* | 1.00000±0.0000 | (Oliveira *et al.*, 2007b) |
| | Egypt | 200 | 0.003 | 0.01674±0.0004* | 1.00000±0.0000 | (Hamdy *et al.*, 2003) |
| | Ghana | 217 | 0 | 0.00371±0.0004* | 0.35011±0.0018 | (Ameyaw *et al.*, 1999) |
| | Mozambique | 250 | 0.002 | 0.00957±0.0005* | 0.53280±0.0020 | (Alves *et al.*, 2004) |
| America | | | | | | |
| | Argentina | 147 | 0.030 | 1.00000±0.0000 | 0.04862±0.0014* | (Laróvere *et al.*, 2003) |
| | Bolivia | 115 | 0.065 | 0.14923±0.0035 | 0.00019±0.0001*† | (Lu *et al.*, 2005) |
| | Brazil | 204 | 0.015 | 0.28452±0.0030 | 0.43680±0.0020 | (Boson *et al.*, 2003) |
| | USA | 282 | 0.032 | 1.00000±0.0000 | 0.01926±0.0010* | (Hon *et al.*, 1999) |
| Asia | | | | | | |
| | China | 192 | 0 | 0.00505±0.0005* | 0.37850±0.0015 | (Collie-Duguid *et al.*, 1999) |
| | Iran | 127 | 0.008 | 0.18815±0.0033 | 1.00000±0.0000 | (Azad *et al.*, 2009) |
| | Japan | 151 | 0 | 0.01018±0.0006* | 0.43493±0.0018 | (Kubota and Chiba, 2001) |
| | Jordania | 169 | 0.012 | 0.24070±0.0037 | 0.65337±0.0024 | (Hakooz *et al.*, 2010) |
| | Korea | 400 | 0 | 0.00036±0.0001*† | 0.22516±0.0024 | (Lee *et al.*, 2008) |
| | North India | 120 | 0.004 | 0.06549±0.0024 | 1.00000±0.0000 | (Kapoor *et al.*, 2009) |
| | South India | 608 | 0 | 0.00016±0.0001*† | 0.15892±0.0014 | (Umamahes waran *et al.*, 2012) |
| | Turkey | 148 | 0.010 | 0.21646±0.0020 | 0.63477±0.0014 | (Sayitoglu *et al.*, 2006) |
| Europe | | | | | | |
| | Croatia | 350 | 0.042 | 0.63597±0.0038 | 0.00410±0.0007* | (Kapitanović *et al.*, 2006) |
| | Denmark | 200 | 0.033 | 1.00000±0.0000 | 0.02198±0.0010* | (Toft *et al.*, 2006) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| | France | 304 | 0.033 | 1.00000±0.0000 | 0.01599±0.0011* | (Ganiere-Monteil *et al.*, 2004) |
| | Germany | 1214 | 0.086 | 0.00953±0.0012* | 0.00000±0.0000*† | (Hakooz *et al.*, 2010) |
| | Italy | 103 | 0.039 | 0.76636±0.0016 | 0.01637±0.0005* | (Rossi *et al.*, 2001) |
| | Norway | 66 | 0.034 | 1.00000±0.0000 | 0.06242±0.0016 | (Hamdy *et al.*, 2003) |
| | Poland | 358 | 0.027 | 0.78055±0.0028 | 0.03703±0.0016* | (Kurzawski *et al.*, 2004) |
| | Portugal | 143 | 0.025 | 0.75567±0.0017 | 0.07861±0.0020 | (Alves *et al.*, 2001) |
| | Russia | 549 | 0.016 | 0.30468±0.0032 | 0.22834±0.0037 | (Samochatova *et al.*, 2009) |
| | Slovenia | 194 | 0.041 | 0.60682±0.0039 | 0.00482±0.0004* | (Hakooz *et al.*, 2010) |
| | Spain | 105 | 0.024 | 1.00000±0.0000 | 0.10468±0.0021 | (Corominas *et al.*, 2003) |
| | Sweden | 800 | 0.038 | 0.81237±0.0027 | 0.00401±0.0007* | (Haglund *et al.*, 2004) |
| | UK | 199 | 0.045 | 0.47493±0.0053 | 0.00268±0.0004* | (Ameyaw *et al.*, 1999) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.002
n, sample size
q, frequencies of the minor allele
Ref., references

Table 10. Comparison of allele frequencies of *TPMT*8* reported from different ethnic populations.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Africa | | | | | | |
| | Cabinda | 103 | 0.024 | 0.08423±0.0017 | 0.10456±0.0021 | (Oliveira *et al.*, 2007b) |
| | Mozambique | 250 | 0.016 | 0.21213±0.0031 | 0.28666±0.0026 | (Alves *et al.*, 2004) |
| America | | | | | | |
| | African-American | 248 | 0.002 | 1.00000±0.0000 | 0.53844±0.0018 | (Hon *et al.*, 1999) |
| Europe | | | | | | |
| | Portugal | 150 | 0.005 | 1.00000±0.0000 | 1.00000±0.0000 | (Alves *et al.*, 2000) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.008
n, sample size
q, frequencies of the minor allele
Ref., references

## 1.4 N-acetyltransferase 2 – NAT2

The NAT2 enzyme that intervenes in many reactions of phase II of the drug metabolism plays a relevant function in the acetylation of several xenobiotics and therapeutic drugs such as is isoniazid (Bakayev *et al.*, 2004; Teixeira *et al.*, 2013).

The presence of polymorphisms in *NAT2* gene can influence the acetylation capacity of the codified enzyme that consequently may alter the activation or detoxification of the genotoxic elements, leading to the manifestation of several side effects (Bakayev *et al.*, 2004).

The *NAT2* 2-SNP genotype panel here used, consisting in the screening of the SNPs c.282C>T and c.341T>C, was reported to allow for the inference of the NAT2 phenotypes with great accuracy, affording a level of precision obtained with the conventional 7-SNP genotype panel. This precision is only not achieved for populations of African ancestry, because this methodology did not include a SNP, c.191G>A, frequently found only in populations from Africa (Cascorbi *et al.*, 1995; Cascorbi and Roots, 1999; Selinski *et al.*, 2011).

Based on these assumptions, the 2-SNP panel was applied to characterise the two Portuguese samples in order to infer the acetylation phenotypes, which are presented in table 11.

Table 11. Frequencies of the slow, intermediate and rapid acetylators in the Portuguese non-Gypsy and Gypsy populations.

| **NAT2 Genotype** | **Portuguese non-Gypsies** | | **Portuguese Gypsies** | |
|---|---|---|---|---|
| | No. of subjects | Frequency | No. of subjects | Frequency |
| 2+ | 40 | 0.571 | 66 | 0.574 |
| 1 | 27 | 0.386 | 43 | 0.374 |
| 0 | 3 | 0.043 | 6 | 0.052 |

The 2-SNP genotype is the sum of variant c.282C>T and c.341T>C alleles with the two and more variant alleles (2+) correspond to the slow genotype, the one variant alleles (1) corresponding to the intermediate genotype and the null variant alleles (0) corresponding to the rapid genotype.

The frequencies of the slow, intermediate or high acetylators (2+, 1 and 0, respectively) were similar in the Portuguese Gypsy and non-Gypsy samples.

The inferred phenotypic data was used to estimate frequencies of the combined set of haplotypes associated to slow and high acetylation capacity, herein referred to for simplicity as slow and high haplotypes, which are shown in table 12. Values in Gypsies and non-Gypsies were quite similar, and so obviously non-significantly different (p>>>>0.05).

Table 12. Combined frequency of slow and rapid *NAT2* haplotypes.

| **Haplotypes** | **Portuguese non-Gypsies** | **Portuguese Gypsies** |
|---|---|---|
| Slow | 0.764±0.036 | 0.761±0.028 |
| Rapid | 0.236±0.036 | 0.239±0.028 |

From previous works that also used the 2-SNP genotype panel to infer the NAT2 phenotypes, information was compiled to carry on comparisons with the haplotypic frequencies obtained in this study (table 13). Up to now, however, data produced with this SNP panel was only available for 5 additional populations, which strongly limits the comparative analysis. From the existing data we can see that the distribution of slow and fast haplotypes is quite homogeneous across populations, including the two population samples here analysed. Consequently, no significant differences were detected between frequencies in the Portuguese Gypsies and non-Gypsies comparatively to any other population.

Table 13. Comparison of slow *NAT2* haplotypes reported from different ethnic populations.

| **Continent** | **Population** | **n** | **Slow** | **P-value Portuguese non-Gypsies** | **P-value Portuguese Gypsies** | **Ref.** |
|---|---|---|---|---|---|---|
| America | | | | | | |
| | Brazil (White Brazilian) | 136 | 0.779 | 0.80865±0.0050 | 0.66750±0.0106 | (Suarez-Kurtz *et al.*, 2012) |
| | Brazil (Guarani) | 88 | 0.722 | 0.43757±0.0069 | 0.42574±0.0124 | (Suarez-Kurtz *et al.*, 2012) |
| | Venezuela | 287 | 0.731 | 0.45752±0.0050 | 0.42480±0.0087 | (Selinski *et al.*, 2011) |
| Asia | | | | | | |
| | Pakistan | 163 | 0.791 | 0.54524±0.0119 | 0.40559±0.0066 | (Selinski *et al.*, 2011) |
| Europe | | | | | | |
| | Hungary plus Germany | 2106 | 0.768 | 0.91989±0.0024 | 0.81453±0.0070 | (Selinski *et al.*, 2011) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.007
n, sample size
Ref., references

## 1.5 Vitamin k epoxide reductase complex subunit 1 – VKORC1

The VKORC1, an important enzyme in the formation of functional clotting factors, is encoded by the recently identified *VKORC1* gene (Li *et al.*, 2004). This enzyme is the target of many anticoagulant or vitamin k antagonists (VKAs) drugs, such as warfarin, inhibiting the synthesis of clotting factors (Yin and Miyata, 2007; D'Andrea *et al.*, 2008; Limdi and Veenstra, 2008).

Many studies have proved that polymorphisms in *VKORC1* gene may influence the sensibility of the individuals to the VKAs, leading to different responses, as bleeding effects or resistance to the therapies. A number of algorithms to adjust the dose administrated to each patient have been proposed taking genetic information into account (Mushiroda *et al.*, 2006; Yin and Miyata, 2007; Puehringer *et al.*, 2010).

The selected SNP that was screened in this work, c.-1639G>A, is one of the most widely studied SNPs in the *VKORC1* gene, due to its well established role in the variability of mRNA level. While the G allele is associated with an increment of 44% in the gene transcription activity, the A allele is responsible for the hindrance of binding of a transcription factor in the VKORC1 promoter region, decreasing the levels of the mature VKORC1 proteins (Sipeky *et al.*, 2009a; Kwon *et al.*, 2011; Smires *et al.*, 2012).

The allele frequency of the *VKORC1* c.-1639G>A in the Portuguese non-Gypsy and Portuguese Gypsy populations are described in Table 14.

Table 14. Allele frequencies of *VKORC1* c.-1639G>A in the Portuguese non-Gypsy and Gypsy populations.

| **VKORC1** | **Portuguese non-Gypsies (n)** | **Portuguese Gypsies (n)** |
|---|---|---|
| c.-1639G>A | 0.443±0.042 (70) | 0.404±0.032 (115) |

n, sample size

The frequency of the derived allele at c.-1639G>A was high in both populations, and although being slightly more elevated in the Portuguese non-Gypsies than in the Gypsies, the difference had no statistical significance (p=0.51784±0.0100).

The estimate now obtained for c.-1639G>A in the reference sample from Portugal is very similar to that previously reported by Jorge *et al*. (2010), which was confirmed when differences between values were tested (Table 15).

When comparing the frequencies obtained in this study with other retrieved from the bibliography (Table 15), we could see that the distribution of c.-1639G>A in the two Portuguese samples was in the range of what has been reported for European populations. Hence, the absence of significant differences between Portuguese Gypsies or non-Gypsies and most other Europeans. The two Portuguese samples also did not differ from the majority of African or American populations. In contrast, compared to populations from Central and Eastern Asia, where allele A reaches frequencies up to 70%, both samples here studied showed highly significant differences.

Respecting to the comparisons with the Hungarian Gypsies studied by Spikey *et al*. (2009a), while both Portuguese Gypsies and non-Gypsies revealed significant differences at the conventional p-value, assuming the p adjusted for multiple tests the difference involving the Portuguese Gypsies lost significance.

Nahar *et al*. (2013) *and* Gaikwad *et al*. (2013), had also examined *VKORC1* c.-1639G>A in North and South India, respectively, reporting frequencies that are in the lower range of values observed at a worldwide scale. Accordingly, when compared to frequencies here estimated for the Portuguese Gypsies and non-Gypsies, all the pairwise differences had statistical significance.

Table 15. Comparison of allele frequencies of *VKORC1* c.-1639G>A reported from different ethnic populations.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Africa | | | | | | |
| | Egypt | 63 | 0.510 | 0.32718±0.0081 | 0.07076±0.0051 | (Bazan *et al.*, 2013) |
| | Morocco | 96 | 0.271 | 0.00228±0.0008* | 0.00611±0.0016* | (Smires *et al.*, 2012) |
| America | | | | | | |
| | African-American | 50 | 0.120 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Wu, 2007) |
| | Argentina | 101 | 0.504 | 0.27048±0.0049 | 0.04594±0.0056* | (Scibona *et al.*, 2012) |
| | Ashkenazi Jewish | 260 | 0.467 | 0.63537±0.0113 | 0.11534±0.0063 | (Scott *et al.*, 2008) |
| | Sephardi Jewish | 80 | 0.500 | 0.35136±0.0060 | 0.06451±0.0068 | (Scott *et al.*, 2008) |
| | USA | 100 | 0.335 | 0.05541±0.0044 | 0.17306±0.0052 | (Babic *et al.*, 2009) |
| Asia | | | | | | |
| | China | 178 | 0.916 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Miao *et al.*, 2007) |
| | Indonesia | 130 | 0.770 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Suriapranata *et al.*, 2011) |
| | Iran | 126 | 0.556 | 0.03737±0.0053* | 0.00110±0.0003*† | (Azarpira *et al.*, 2010) |
| | Japan | 341 | 0.918 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Yoshizawa *et al.*, 2009) |
| | Korea | 37 | 0.945 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Kwon *et al.*, 2011) |
| | Lebanon | 231 | 0.524 | 0.10070±0.0075 | 0.00445±0.0013* | (Esmerian *et al.*, 2011) |
| | Malaysia | 91 | 0.736 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Gan *et al.*, 2011) |
| | North India | 209 | 0.190 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Nahar *et al.*, 2013) |
| | Saudi Arabia | 131 | 0.427 | 0.83956±0.0024 | 0.65138±0.0079 | (Alzahrani *et al.*, 2013) |
| | South India | 145 | 0.128 | 0.00000±0.0000*† | 0.00000±0.0000*† | (Gaikwad *et al.*, 2013) |

Continuation of the previous table.

| Continent | Population | n | q | P-value Portuguese non-Gypsies | P-value Portuguese Gypsies | Ref. |
|---|---|---|---|---|---|---|
| Europe | Turkey | 292 | 0.404 | 0.44316±0.0097 | 1.00000±0.0000 | (Silan *et al.*, 2012) |
| | France | 263 | 0.420 | 0.63444±0.0068 | 0.73965±0.0122 | (Bodin *et al.*, 2005) |
| | Hungary | 510 | 0.390 | 0.23282±0.0048 | 0.70275±0.0040 | (Sipeky *et al.*, 2009a) |
| | Hungary (Gypsy) | 451 | 0.297 | 0.00083±0.0001*† | 0.00237±0.0003* | (Sipeky *et al.*, 2009a) |
| | Italy | 437 | 0.443 | 1.00000±0.0000 | 0.32704±0.0174 | (Zambon *et al.*, 2010) |
| | Portugal | 91 | 0.417 | 0.65522±0.0075 | 0.84137±0.0046 | (Jorge *et al.*, 2010) |
| | Romania | 332 | 0.422 | 0.71697±0.0087 | 0.69704±0.0078 | (Buzoianu *et al.*, 2012) |
| | Russia | 111 | 0.333 | 0.04773±0.0043* | 0.12143±0.0087 | (Vorob'eva *et al.*, 2011) |
| | Spain | 101 | 0.436 | 0.91115±0.0026 | 0.54417±0.0070 | (Scott *et al.*, 2010) |
| | Sweden | 1461 | 0.394 | 0.25094±0.0056 | 0.78772±0.0058 | (Wadelius *et al.*, 2009) |

* Significant value at p≤0.05
† Significant value with Bonferroni correction p<0.002
n, sample size
q, frequencies of the minor allele
Ref., references

Highlighting the frequencies for *VKORC1* c.-1639G>A in European, Portuguese Gypsies and Indian populations, the pattern of differences is consistent, although less sharp, with that observed for *CYP2C9*2* and *CYP2C19*2*, indicating that the evolutionary history of a population can influence differently each gene (Figure 12).

Figure 12. *VKORC1* c.-1639G>A allele frequency distribution of the European, Indian and Gypsy populations. The frequency of the Portuguese Gypsy population is represented in blue, the frequency of the Portuguese non-Gypsy population is represented in red and the frequency of the Hungarian Roma population is represented in violet.

## 2. Final considerations

The interindividual and interethnic variability in the response to drugs can be influenced by polymorphisms in genes encoding metabolising enzymes (DMEs), transporters and targets of the drugs, altering their normal metabolism and effect. Thus, the development of the Pharmacogenetic tests can be a helpful tool in the clinical practice, because the mobilization of relevant information about the genes and polymorphisms that affect the metabolism and effectiveness of drugs, can improve the therapeutic protocols.

In the present study, we have analysed variants in 4 genes coding for distinct DMEs and 1 gene that encodes a drug target in Portuguese Gypsy and non-Gypsy samples.

The results obtained, only revealed significant differences between both samples in one of the screened variations, which was *CYP2C9*2* that occurred at frequency of 10.8% in Gypsies and 20.0% in non-Gypsies. Within the same locus, the frequency of *CYP2C9*3* was, however, non-significantly higher in Gypsies than in non-Gypsies, implying that for *CYP2C9*, the combined frequency of alleles associated to decreased enzymatic activity attained 18.7% and 25.7% in the Portuguese Gypsies and non-Gypsies, respectively. This means that in Portugal, Gypsies, comparatively to other Portuguese, are slightly less predisposed to develop ADRs for drugs metabolised by CYP2C9. Applying the Fisher exact test, the differences observed between the combined frequency of alleles in Portuguese Gypsies and non-Gypsies had no statistical significance ($p= 0.14599\pm0.0053$).

For the other studied pharmacogenetic polymorphisms, no significant differences were detected between Gypsies and other Portuguese.

As a whole, our findings indicate that from the Pharmacogenetic point of view, the Gypsies from Portugal do not raise added concerns comparatively to the host population.

The analysis of pharmacogenetic-related diversity in the Portuguese Gypsies was triggered by the previous knowledge that they showed peculiar genetic characteristics that distinguish them, as a group, from the Portuguese, in the general sense. Furthermore, all Gypsy groups have been subject along history to strong drift effects, as is for instance reflected in their fraction of genetic diversity that implies pathogenic consequences. In fact, many diseases occur with unusual high prevalence in Roma groups, as exemplifies the elevated incidence of Maple Syrup Urine Disease in the Portuguese Gypsies (Quental *et al.*, 2009).

This justifies the rationale of investigating in the Portuguese Gypsies genetic variations known to have pharmacogenetic relevance. For none of the polymorphisms examined, the Portuguese Gypsies, comparatively to other Portuguese, appear to need a special vigilance care, in the prescription, administration, dosage and maintenance of metabolised drugs by the enzymes encoded by the genes studied.

Obviously, the picture here obtained on the pharmacogenetic profile of the Portuguese Gypsies still needs to be improved.

A critical drawback of this study was to be based on a small number of SNPs for each gene, which despite having been selected among the most widely studied SNPs, may not represent the adequate SNP coverage for Gypsy groups. In fact, without knowing at which extent a given set of screened variations contribute to explain variability in levels of enzymatic/proteic activity in a population, it is difficult to evaluate whether it represent a good proxy of the set of alleles affecting gene expression in that population. In the case of Gypsies, this might be a key question, due to their shared past ancestry in Indian populations, which currently also remain understudied in which respect pharmacogenetic polymorphisms. By other words, we cannot exclude that, for each of the studied genes, additional variations apart from those examined might impact substantially drug response in Gypsies, whilst being irrelevant (due to absence or rarity) in other European populations.

To clarify the issue, enzymatic/proteic activities should be measured in individuals from the two populations and then contrasted with the genotypic results obtained through a fine gene screening. This would elucidate whether unknown or unstudied SNPs have pharmacogenetic consequences in the Portuguese Gypsies, which may be rare or absent in the general Portuguese population.

If the screened variations only encompassed a very small fraction of pharmacogenetic polymorphisms, it is noteworthy how they captured signals of the history of Portuguese Gypsies. Most of the variations had frequency data for Indian populations, which showed values rather different from those typically found in European populations, including from Portugal. For all these variation the Portuguese Gyspies presented frequency values between those found in Indians and Portuguese non-Gypsies. Despite the absence of a systematic statistical support of the differences, the finding is fairly consistent with the inferred genetic origin of the Roma, since strong evidence point to India as the region from where departed the first people who, in successive waves of migrations, would give rise to the European Roma.

In Figure 13 is depicted a plot obtained through Principal Component Analysis based on frequencies of *CYP2C9*2*, *CYP2C9*3, CYP2C19*2*, *TPMT*3A* and *VKORC1* c.-1639G>A in distinct European and Indian populations. Only samples with data

available for the five variations were included in the analysis, explaining the relative small number of populations considered. *NAT2* and *TPMT*8* was excluded because it had not been characterized in those populations.

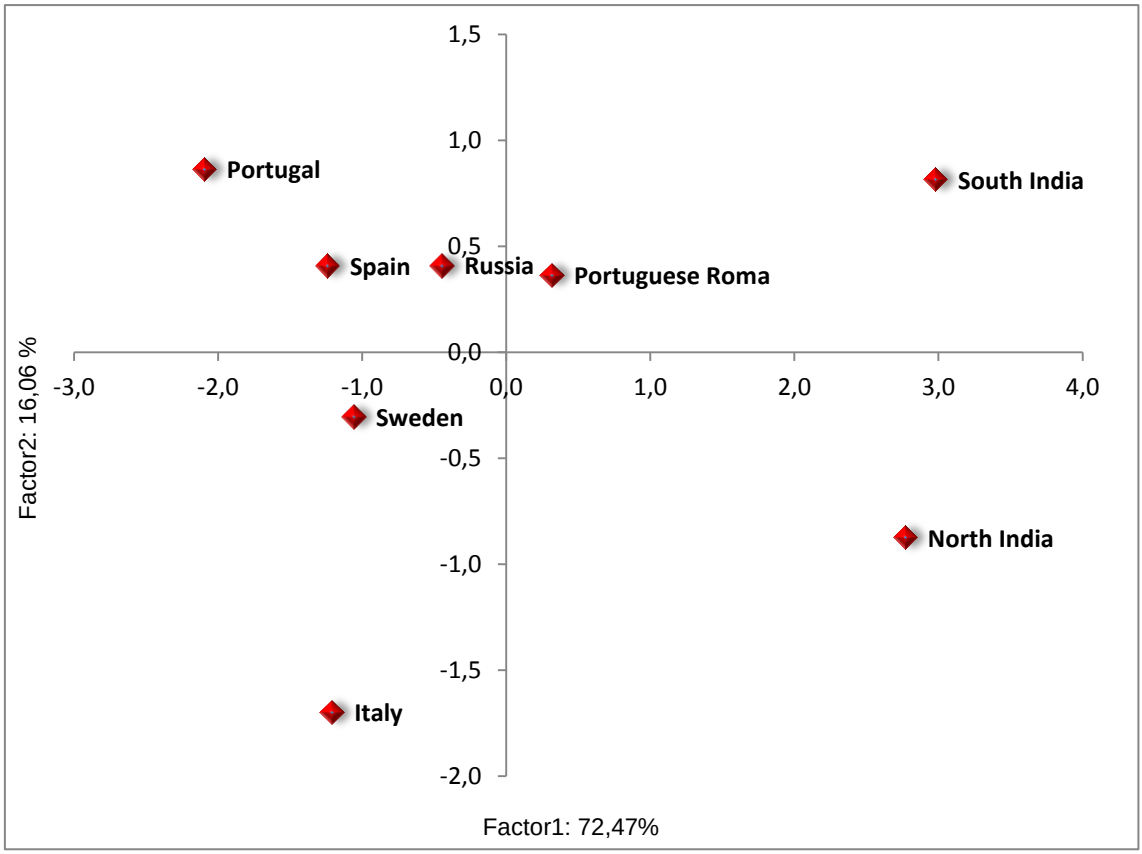


Figure 13. Principal Component Analysis (PCA) based on allelic frequency values between 8 populations (Portugal, Spain, Russia, Portuguese Roma, Sweden, Italy, South India, North India) for 5 SNPs.

Dimension 1 sharply discriminates Indian from European populations. The Gypsies from Portugal are integrated in the cluster of European populations, but among them the Gypsies are the group positioned closer to Indian samples, seemly testifying that Indians were one of their source populations.

# V. CONCLUSIONS

Many populations of the world have been the subject of the pharmacogenetic studies, however there are still population groups poorly studied in this area, as is the case of the Gypsy groups.

Gypsies meet unique genetic characteristics that differentiate them from the host populations, making them a very attractive case of study in several areas.

In this work, several SNPs in genes with pharmacogenetic relevance were analysed in a sample of Portuguese Gypsies. Then, the results were compared with those respecting to the host population, in order to evaluate whether significant differences between the populations existed.

The obtained results did not demonstrate major differences between the two populations for the studied SNPs. Thus, this means that it is not necessary a distinct vigilance care in the Portuguese Gypsies in comparison with other Portuguese in the administration of drugs whose metabolism or effect depend on the encoded products of the studied genes.

The absence of significant differences between the two populations do not necessarily signifies that they have similar pharmacogenetic risks, since it was only analysed a small group of genes and SNPs, which may not be enough to bring to light unsuspected differences in terms of the expectations on how individuals from each population will respond to drugs.

# VI. FUTURE DIRECTIONS

In order to obtain a better assessment of differences between Portuguese Gypsies and the host population, it should be quantified the activity of the enzymes/proteins encoded by the selected genes and contrast the information with genotypic data. Preferably complete sequencing of the genes should be performed, which would allow for the detection of new SNPs in the Portuguese Gypsies that might be of pharmacogenetical significance. In addition, other genes and SNPs known to affect drug response should be studied and compared in both samples.

It would be also important to investigate thoroughly other Gypsy groups from all around the world, thus allowing a better understanding of their pharmacogenetic profiles, once none Gypsy group can be representative of other Gypsies groups.

# VII. BIBLIOGRAPHY

Allabi, A.C., Gala, J.L., Desager, J.P., Heusterspreute, M., and Horsmans, Y. (2003). Genetic polymorphisms of *CYP2C9* and *CYP2C19* in the Beninese and Belgian populations. Br. J. Clin. Pharmacol. *56*, 653-657.

Alves, S., Amorim, A., Ferreira, F., and Prata, M.J. (2001). Influence of the variable number of tandem repeats located in the promoter region of the thiopurine methyltransferase gene on enzymatic activity. Clin. Pharmacol. Ther. *70*, 165-174.

Alves, S., Prata, M.J., Ferreira, F., and Amorim, A. (2000). Screening of thiopurine S-methyltransferase mutations by horizontal conformation-sensitive gel electrophoresis. Hum. Mutat. *15*, 246-253.

Alves, S., Rocha, J., Amorim, A., and Prata, M.J. (2004). Tracing the Origin of the Most Common Thiopurine Methyltransferase (*TPMT*) Variants: Preliminary Data from the Patterns of Haplotypic Association with Two CA Repeats. Ann. Hum. Genet. *68*, 313-323.

Alzahrani, A.M., Ragia, G., Hanieh, H., and Manolopoulos, V.G. (2013). Genotyping of *CYP2C9* and *VKORC1* in the Arabic Population of Al-Ahsa, Saudi Arabia. Biomed Res. Int. *2013*, 6.

Ameyaw, M.M., Collie-Duguid, E.S., Powrie, R.H., Ofori-Adjei, D., and McLeod, H.L. (1999). Thiopurine Methyltransferase Alleles in British and Ghanaian Populations. Hum. Mol. Genet. *8*, 367-370.

Arias, B., Catalán, R., Gastó, C., Gutiérrez, B., and Fañanás, L. (2005). Evidence for a combined genetic effect of the 5-HT1A receptor and serotonin transporter genes in the clinical outcome of major depressive patients treated with citalopram. J. Psychopharmacol. *19*, 166-172.

Arvanitidis, K., Ragia, G., Iordanidou, M., Kyriaki, S., Xanthi, A., Tavridou, A., and Manolopoulos, V.G. (2007). Genetic polymorphisms of drug-metabolizing enzymes CYP2D6, CYP2C9, CYP2C19 and CYP3A5 in the Greek population. Fundam. Clin. Pharmacol. *21*, 419-426.

Azad, M., Kaviani, S., Soleimani, M., Noruzinia, M., and A., H. (2009). Common Polymorphism's Analysis of Thiopurine S-Methyltransferase (*TPMT*) in Iranian Population. Yakhteh. Med. J. *11*, 311-316.

Azarpira, N., Namazi, S., Hendijani, F., Banan, M., and Darai, M. (2010). Investigation of allele and genotype frequencies of *CYP2C9*, *CYP2C19* and *VKORC1* in Iran. Pharmacol. Rep. *62*, 740-746.

Babaoglu, M.O., Yasar, U., Sandberg, M., Eliasson, E., Dahl, M.L., Kayaalp, S.O., and Bozkurt, A. (2004). *CYP2C9* genetic variants and losartan oxidation in a Turkish population. Eur. J. Clin. Pharmacol. *60*, 337-342.

Babic, N., Haverfield, E.V., Burrus, J.A., Lozada, A., Das, S., and Yeo, K.T. (2009). Comparison of performance of three commercial platforms for warfarin sensitivity genotyping. Clin. Chim. Acta *406*, 143-147.

Bae, J.W., Kim, H.K., Kim, J.H., Yang, S.I., Kim, M.J., Jang, C.G., Park, Y.S., and Lee, S.Y. (2005). Allele and genotype frequencies of *CYP2C9* in a Korean population. Br. J. Clin. Pharmacol. *60*, 418-422.

Bakayev, V.V., Mohammadi, F., Bahadori, M., Sheikholslami, M., Javeri, A., Masjedi, M.R., and Velayati, A.A. (2004). Arylamine N-acetyltransferase 2 slow acetylator polymorphisms in unrelated Iranian individuals. Eur. J. Clin. Pharmacol. *60*, 467-471.

Bathum, L., Andersen-Ranberg, K., Boldsen, J., Brøsen, K., and Jeune, B. (1998). Genotypes for the cytochrome P450 enzymes CYP2D6 and CYP2C19 in human longevity Role of CYP2D6 and CYP2C19 in longevity. Eur. J. Clin. Pharmacol. *54*, 427-430.

Bazan, N.S., Sabry, N.A., Rizk, A., Mokhtar, S., and Badary, O.A. (2013). Factors affecting warfarin dose requirements and quality of anticoagulation in adult Egyptian patients: role of gene polymorphism. Ir. J. Med. Sci., 1-12.

Bodin, L., Verstuyft, C., Tregouet, D.A., Robert, A., Dubert, L., Funck-Brentano, C., Jaillon, P., Beaune, P., Laurent-Puig, P., Becquemont, L.*, et al.* (2005). Cytochrome P450 2C9 (*CYP2C9*) and vitamin K epoxide reductase (*VKORC1*) genotypes as determinants of acenocoumarol sensitivity. Blood *106*, 135-140.

Bonello, L., Tantry, U.S., Marcucci, R., Blindt, R., Angiolillo, D.J., Becker, R., Bhatt, D.L., Cattaneo, M., Collet, J.P., Cuisset, T.*, et al.* (2010). Consensus and Future Directions on the Definition of High On-Treatment Platelet Reactivity to Adenosine Diphosphate. J. Am. Coll. Cardiol. *56*, 919-933.

Boson, W.L., Romano-Silva, M.A., Correa, H., Falcao, R.P., Teixeira-Vidigal, P.V., and De Marco, L. (2003). Thiopurine methyltransferase polymorphisms in a Brazilian population. Pharmacogenomics J. *3*, 178-182.

Bozina, N., Granić, P., Lalić, Z., Tramisak, I., Lovrić, M., and Stavljenić-Rukavina, A. (2003). Genetic polymorphisms of cytochromes P450: *CYP2C9*, *CYP2C19*, and *CYP2D6* in Croatian population. Croat. Med. J. *44*, 425-428.

Bravo-Villalta, H., Yamamoto, K., Nakamura, K., Bayá, A., Okada, Y., and Horiuchi, R. (2005). Genetic polymorphism of *CYP2C9* and *CYP2C19* in a Bolivian population: an investigative and comparative study. Eur. J. Clin. Pharmacol. *61*, 179-184.

Brockmöller, J., and Tzvetkov, M. (2008). Pharmacogenetics: data, concepts and tools to improve drug discovery and drug treatment. Eur. J. Clin. Pharmacol. *64*, 133-157.

Burian, M., Grösch, S., Tegeder, I., and Geisslinger, G. (2002). Validation of a new fluorogenic real-time PCR assay for detection of *CYP2C9* allelic variants and *CYP2C9* allelic distribution in a German population. Br. J. Clin. Pharmacol. *54*, 518-521.

Buzoianu, A.D., Trifa, A.P., Mureşanu, D.F., and Crişan, S. (2012). Analysis of *CYP2C9*2, CYP2C9*3* and *VKORC1* -1639 G>A polymorphisms in a population from South-Eastern Europe. J. Cell. Mol. Med. *16*, 2919-2924.

Buzoianu, A.D., Trifa, A.P., Popp, R.A., Militaru, M.S., Militaru, C.F., Bocşan, C.I., Farcaş, M.F., and Pop, I.V. (2010). Screening for *CYP2C19*2*, **3* and **4* gene variants in a Romanian population study group. Farmacia *58*, 806-817.

Carlquist, J.F., Horne, B.D., Muhlestein, J.B., Lappé, D.L., Whiting, B.M., Kolek, M.J., Clarke, J.L., James, B.C., and Anderson, J.L. (2006). Genotypes of the cytochrome p450 isoform, *CYP2C9*, and the vitamin K epoxide reductase complex subunit 1 conjointly determine stable warfarin dose: a prospective study. J. Thromb. Thrombolysis *22*, 191-197.

Cascorbi, I., Drakoulis, N., Brockmöller J, Maurer, A., Sperling, K., and Roots, I. (1995). Arylamine N-acetyltransferase (*NAT2*) mutations and their allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity. Am. J. Hum. Genet. *57*, 581-592.

Cascorbi, I., and Roots, I. (1999). Pitfalls in N-acetyltransferase 2 genotyping. Pharmacogenetics *9*, 123-127.

Cavallari, L.H., Jeong, H., and Bress, A. (2011). Role of cytochrome *P450* genotype in the steps toward personalized drug therapy. Pharmgenomics Pers. Med. *4*, 123-136.

Chaudhry, A.S., Kochhar, R., and Kohli, K.K. (2008). Genetic polymorphism of *CYP2C19* & therapeutic response to proton pump inhibitors. Indian J. Med. Res. *127*, 521-530.

Collie-Duguid, E.S., Pritchard, S.C., Powrie, R.H., Sludden, J., Collier, D.A., Li, T., and McLeod, H.L. (1999). The frequency and distribution of thiopurine methyltransferase alleles in Caucasian and Asian populations. Pharmacogenetics *9*, 37-42.

Corominas, H., Domènech, M., Laíz, A., Gich, I., Geli, C., Díaz, C., de Cuevillas, F., Moreno, M., Vázquez, G., and Baiget, M. (2003). Is thiopurine methyltransferase genetic polymorphism a major factor for withdrawal of azathioprine in rheumatoid arthritis patients? Rheumatology (Oxford) *42*, 40-45.

Crews, K.R., Hicks, J.K., Pui, C.H., Relling, M.V., and Evans, W.E. (2012). Pharmacogenomics and Individualized Medicine: Translating Science Into Practice. Clin. Pharmacol. Ther. *92*, 467-475.

D'Andrea, G., D'Ambrosio, R., and Margaglione, M. (2008). Oral anticoagulants: Pharmacogenetics: Relationship between genetic and non-genetic factors. Blood Rev. *22*, 127-140.

Daly, A.K. (2010). Pharmacogenetics and human genetic polymorphisms. Biochem. J. *429*, 435-449.

Dandara, C., Lombard, Z., Du Plooy, I., McLellan, T., Norris, S.A., and Ramsay, M. (2011). Genetic variants in *CYP* (*-1A2*, *-2C9*, *-2C19*, *-3A4* and *-3A5*), *VKORC1* and *ABCB1* genes in a black South African population: a window into diversity. Pharmacogenomics *12*, 1663-1670.

Danielson, P.B. (2002). The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. Curr. Drug Metab. *3*, 561-597.

Dickmann, L.J., Rettie, A.E., Kneller, M.B., Kim, R.B., Wood, A.J., Stein, C.M., Wilkinson, G.R., and Schwarz, U.I. (2001). Identification and Functional Characterization of a New *CYP2C9* Variant (*CYP2C9*5*) Expressed among African Americans. Mol. Pharmacol. *60*, 382-387.

Djaffar Jureidini, I., Chamseddine, N., Keleshian, S., Naoufal, R., Zahed, L., and Hakime, N. (2011). Prevalence of *CYP2C19* polymorphisms in the Lebanese population. Mol. Biol. Rep. *38*, 5449-5452.

Esmerian, M.O., Mitri, Z., Habbal, M.Z., Geryess, E., Zaatari, G., Alam, S., Skouri, H.N., Mahfouz, R.A., Taher, A., and Zgheib, N.K. (2011). Influence of *CYP2C9* and *VKORC1* Polymorphisms on Warfarin and Acenocoumarol in a Sample of Lebanese People. J. Clin. Pharmacol. *51*, 1418-1428.

Evans, W.E., and McLeod, H.L. (2003). Pharmacogenomics — Drug Disposition, Drug Targets, and Side Effects. N. Engl. J. Med. *348*, 538-549.

Excoffier, L., and Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. *10*, 564-567.

Fehér, Á., Juhász, A., Rimanóczy, Á., Álmos, P., Béres, J., Janka, Z., and Kálmán, J. (2011). Dopamine metabolism-related gene polymorphisms in Roma (Gypsy) and Hungarian populations. J. Genet. *90*, e72-e75.

Fishbain, D.A., Fishbain, D., Lewis, J., Cutler, R.B., Cole, B., Rosomoff, H.L., and Rosomoff, R.S. (2004). Genetic Testing for Enzymes of Drug Metabolism: Does It Have Clinical Utility for Pain Medicine at the Present Time? A Structured Review. Pain Med. *5*, 81-93.

Furuta, T., Shirai, N., Sugimoto, M., Nakamura, A., Hishida, A., and Ishizaki, T. (2005). Influence of *CYP2C19* Pharmacogenetic Polymorphism on Proton Pump Inhibitor-based Therapies. Drug Metab. Pharmacokinet. *20*, 153-167.

Fuselli, S., Gilman, R.H., Chanock, S.J., Bonatto, S.L., De Stefano, G., Evans, C.A., Labuda, D., Luiselli, D., Salzano, F.M., Soto, G.*, et al.* (2007). Analysis of nucleotide diversity of *NAT2* coding region reveals homogeneity across Native American populations and high intra-population diversity. Pharmacogenomics J. *7*, 144-152.

Gaikovitch, E., Cascorbi, I., Mrozikiewicz, P., Brockmoller, J., Frotschl, R., Kopke, K., Gerloff, T., Chernov, J., and Roots, I. (2003). Polymorphisms of drug-metabolizing enzymes CYP2C9, CYP2C19, CYP2D6, CYP1A1, NAT2 and of P-glycoprotein in a Russian population. Eur. J. Clin. Pharmacol. *59*, 303 - 312.

Gaikwad, T., Ghosh, K., Kulkarni, B., Kulkarni, V., Ross, C., and Shetty, S. (2013). Influence of CYP2C9 and VKORC1 gene polymorphisms on warfarin dosage, over anticoagulation and other adverse outcomes in Indian population. Eur. J. Pharmacol. *710*, 80-84.

Gan, G., Phipps, M., Lee, M.T., Lu, L., Subramaniam, R., Bee, P., and Chang, S. (2011). Contribution of VKORC1 and CYP2C9 polymorphisms in the interethnic variability of warfarin dose in Malaysian populations. Ann. Hematol. *90*, 635-641.

Ganiere-Monteil, C., Medard, Y., Lejus, C., Bruneau, B., Pineau, A., Fenneteau, O., Bourin, M., and Jacqz-Aigrain, E. (2004). Phenotype and genotype for thiopurine methyltransferase activity in the French Caucasian population: impact of age. Eur. J. Clin. Pharmacol. *60*, 89-96.

Garat, A., Cauffiez, C., Renault, N., Lo-Guidice, J.M., Allorge, D., Chevalier, D., Houdret, N., Chavatte, P., Loriot, M.A., Gala, J.L.*, et al.* (2008). Characterisation of novel defective thiopurine S-methyltransferase allelic variants. Biochem. Pharmacol. *76*, 404-415.

Gawrońska-Szklarz, B., Siuda, A., Kurzawski, M., Bielicki, D., Marlicz, W., and Drożdzik, M. (2010). Effects of *CYP2C19*, *MDR1*, and interleukin *1-B* gene variants on the eradication rate of *Helicobacter pylori* infection by triple therapy with pantoprazole, amoxicillin, and metronidazole. Eur. J. Clin. Pharmacol. *66*, 681-687.

Gonzalez, F.J., and Turkey, R.H. (2006). Drug Metabolism. In Goodman and Gilman's pharmacological basis of therapeutics, L.L. Brunton, J.S. Lazo, and K.L. Parker, eds. (New York: McGraw-Hill Medical Publishing Division), pp. 71-92.

Gresham, D., Morar, B., Underhill, P.A., Passarino, G., Lin, A.A., Wise, C., Angelicheva, D., Calafell, F., Oefner, P.J., Shen, P.*, et al.* (2001). Origins and Divergence of the Roma (Gypsies). Am. J. Hum. Genet. *69*, 1314-1331.

Guo, L., Dial, S., Shi, L., Branham, W., Liu, J., Fang, J.L., Green, B., Deng, H., Kaput, J., and Ning, B. (2011). Similarities and Differences in the Expression of Drug-Metabolizing Enzymes between Human Hepatic Cell Lines and Primary Human Hepatocytes. Drug Metab. Dispos. *39*, 528-538.

Gusmão, A., Gusmão, L., Gomes, V., Alves, C., Calafell, F., Amorim, A., and Prata, M.J. (2008). A Perspective on the History of the Iberian Gypsies Provided by Phylogeographic Analysis of Y-Chromosome Lineages. Ann. Hum. Genet. *72*, 215-227.

Gusmão, A., Valente, C., Gomes, V., Alves, C., Amorim, A., Prata, M.J., and Gusmão, L. (2010). A genetic historical sketch of European Gypsies: The perspective from autosomal markers. Am. J. Phys. Anthropol. *141*, 507-514.

Haglund, S., Lindqvist, M., Almer, S., Peterson, C., and Taipalensuu, J. (2004). Pyrosequencing of *TPMT* Alleles in a General Swedish Population and in Patients with Inflammatory Bowel Disease. Clin. Chem. *50*, 288-295.

Hakooz, N., Arafat, T., Payne, D., Ollier, W., Pushpakom, S., Andrews, J., and Newman, W. (2010). Genetic analysis of thiopurine methyltransferase polymorphism in the Jordanian population. Eur. J. Clin. Pharmacol. *66*, 999-1003.

Hamdy, S.I., Hiratsuka, M., Narahara, K., El-Enany, M., Moursi, N., Ahmed, M.S., and Mizugaki, M. (2002). Allele and genotype frequencies of polymorphic cytochromes P450 (*CYP2C9, CYP2C19, CYP2E1*) and dihydropyrimidine dehydrogenase (*DPYD*) in the Egyptian population. Br. J. Clin. Pharmacol. *53*, 596-603.

Hamdy, S.I., Hiratsuka, M., Narahara, K., Endo, N., El-Enany, M., Moursi, N., Ahmed, M.S.E., and Mizugaki, M. (2003). Genotype and allele frequencies of *TPMT*, *NAT2*, *GST*, *SULT1A1* and *MDR-1* in the Egyptian population. Br. J. Clin. Pharmacol. *55*, 560-569.

Hein, D.W., and Doll, M.A. (2012). Accuracy of various human *NAT2* SNP genotyping panels to infer rapid, intermediate and slow acetylator phenotypes. Pharmacogenomics *13*, 31-41.

Herman, D., Dolžan, V., and Breskvar, K. (2003). Genetic polymorphism of cytochromes P450 2C9 and 2C19 in Slovenian population. Zdrav. Vestn. *72*, 347-351.

Hiratsuka, M., Sasaki, T., and Mizugaki, M. (2006). Genetic testing for pharmacogenetics and its clinical application in drug therapy. Clin. Chim. Acta *363*, 177-186.

Holmes, D.R.J., Dehmer, G.J., Kaul, S., Leifer, D., O'Gara, P.T., and Stein, C.M. (2010). ACCF/AHA Clopidogrel Clinical Alert: Approaches to the FDA “Boxed Warning”A Report of the American College of Cardiology Foundation Task Force on Clinical Expert Consensus Documents and the American Heart Association Endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. J. Am. Coll. Cardiol. *56*, 321-341.

Hon, Y.Y., Fessing, M.Y., Pui, C.H., Relling, M.V., Krynetski, E.Y., and Evans, W.E. (1999). Polymorphism of the Thiopurine S-Methyltransferase Gene in African-Americans. Hum. Mol. Genet. *8*, 371-376.

Hoskins, J.M., Shenfield, G.M., and Gross, A.S. (1998). Relationship between proguanil metabolic ratio and *CYP2C19* genotype in a Caucasian population. Br. J. Clin. Pharmacol. *46*, 499-504.

Hunfeld, N.G., Mathot, R.A., Touw, D.J., Van Schaik, R.H., Mulder, P.G., Franck, P.F., Kuipers, E.J., and Geus, W.P. (2008). Effect of *CYP2C19*2* and **17* mutations on pharmacodynamics and kinetics of proton pump inhibitors in Caucasians. Br. J. Clin. Pharmacol. *65*, 752-760.

Imming, P., Sinning, C., and Meyer, A. (2006). Drugs, their targets and the nature and number of drug targets. Nat. Rev. Drug Discov. *5*, 821-834.

Ingelman-Sundberg, M., Sim, S.C., Gomez, A., and Rodriguez-Antona, C. (2007). Influence of cytochrome P450 polymorphisms on drug therapies: Pharmacogenetic, pharmacoepigenetic and clinical aspects. Pharmacol. Ther. *116*, 496-526.

Isaza, C., Henao, J., Martínez, J.H., Sepúlveda Arias, J.C., and Beltrán, L. (2007). Phenotype-genotype analysis of *CYP2C19* in Colombian mestizo individuals. BMC Clin. Pharmacol. *7*, 6.

Johnson, J.A. (2003). Pharmacogenetics: potential for individualized drug therapy through genetics. Trends Genet. *19*, 660-666.

Jorge, E., Pêgo, J.M., Baptista, R., Lourenço, M., Marques, I., Correia, A., Monteiro, P., Antunes, F., and Providência, L.A. (2010). Genetic characterization of warfarin sensitivity in a population of cardiovascular patients on chronic anticoagulation. Rev. Port. Cardiol. *29*, 1831-1838.

Jose, R., Chandrasekaran, A., Sam, S.S., Gerard, N., Chanolean, S., Abraham, B.K., Satyanarayanamoorthy, K., Peter, A., and Rajagopal, K. (2005). *CYP2C9* and *CYP2C19* genetic polymorphisms: frequencies in the south Indian population. Fundam. Clin. Pharmacol. *19*, 101-105.

Kalaydjieva, L., Gresham, D., and Calafell, F. (2001). Genetic studies of the Roma (Gypsies): a review. BMC Med. Genet. *2*, 5.

Kalaydjieva, L., Morar, B., Chaix, R., and Tang, H. (2005). A newly discovered founder population: the Roma/Gypsies. Bioessays *27*, 1084-1094.

Kalow, W. (2002). Pharmacogenetics and personalised medicine. Fundam. Clin. Pharmacol. *16*, 337-342.

Kalow, W. (2006). Pharmacogenetics and pharmacogenomics: origin, status, and the hope for personalized medicine. Pharmacogenomics J. *6*, 162–165.

Kapitanović, S., Jokić, M., and Jurišić, G. (2006). *TPMT* Gene Polymorphisms in Croatian Population. In European Human Genetics Conference 2006 (Amsterdam Nature Publishing Group), pp. 353-353.

Kapoor, G., Maitra A., Somlata, and Brahmachari, V. (2009). Application of SNaPshot for analysis of thiopurine methyltransferase gene polymorphism. Indian J. Med. Res. *129*, 500-505.

Kim, K.A., Song, W.K., Kim, K.R., and Park, J.Y. (2010). Assessment of *CYP2C19* genetic polymorphisms in a korean population using a simultaneous multiplex pyrosequencing method to simultaneously detect the *CYP2C19*2*, *CYP2C19*3*, and *CYP2C19*17* alleles. J. Clin. Pharm. Ther. *35*, 697-703.

Krynetski, E., and Evans, W.E. (2003). Drug methylation in cancer therapy: lessons from the *TPMT* polymorphism. Oncogene *22*, 7403-7413.

Kubota, T., and Chiba, K. (2001). Frequencies of thiopurine S-methyltransferase mutant alleles (*TPMT*2, *3A, *3B* and **3C*) in 151 healthy Japanese subjects and the inheritance of *TPMT*3C* in the family of a propositus. Br. J. Clin. Pharmacol. *51*, 475–477.

Kudzi, W., Adjei, G.O., Ofori-Adjei, D., and Dodoo, A.N. (2011). Pharmacogenetics in Ghana: Reviewing the evidence. Ghana Med. J. *45*, 73-80.

Kudzi, W., Dodoo, A.N., and Mills, J.J. (2009). Characterisation of *CYP2C8*, *CYP2C9* and *CYP2C19* polymorphisms in a Ghanaian population. BMC Med. Genet. *10*, 124.

Kupiec, T.C., Raj, V., and Vu, N. (2006). Pharmacogenomics for the Forensic Toxicologist. J. Anal. Toxicol. *30*, 65-72.

Kurzawski, M., Gawronska-Szklarz, B., and Drozdzik, M. (2004). Frequency Distribution of Thiopurine S-Methyltransferase Alleles in a Polish Population. Ther. Drug Monit. *26*, 541-545.

Kwon, A., Jo, S.H., Im, H.J., Jo, Y.A., Park, J.Y., Kang, H.J., Kim, H.S., Cho, H.C., and Lee, Y.K. (2011). Pharmacogenetic distribution of warfarin and its clinical significance in Korean patients during initial anticoagulation therapy. J. Thromb. Thrombolysis *32*, 467-473.

Lam, M.P., and Cheung, B.M. (2012). The pharmacogenetics of the response to warfarin in Chinese. Br. J. Clin. Pharmacol. *73*, 340-347.

Lanfear, D.E., and McLeod, H.L. (2007). Pharmacogenetics: Using DNA to Optimize Drug Therapy. Am. Fam. Physician. *76*, 1179-1182.

Lareu, M.V., Phillips, C.P., Carracedo, A., Lincoln, P.J., Syndercombe Court, D., and Thomson, J.A. (1994). Investigation of the STR locus *HUMTH01* using PCR and two electrophoresis formats: UK and Galician Caucasian population surveys and usefulness in paternity investigations. Forensic Sci. Int. *66*, 41-52.

Laróvere, L.E., de Kremer, R.D., Lambooy, L.H., and De Abreu, R.A. (2003). Genetic polymorphism of thiopurine S-methyltransferase in Argentina. Ann. Clin. Biochem. *40*, 388-393.

Lee, S.J., Kim, W.Y., Kim, H., Shon, J.H., Lee, S.S., and Shin, J.G. (2009). Identification of New *CYP2C19* Variants Exhibiting Decreased Enzyme Activity in the Metabolism of S-Mephenytoin and Omeprazole. Drug Metab. Dispos. *37*, 2262-2269.

Lee, S.S., Kim, W.Y., Jang, Y.J., and Shin, J.G. (2008). Duplex pyrosequencing of the *TPMT*3C* and *TPMT*6* alleles in Korean and Vietnamese populations. Clin. Chim. Acta *398*, 82-85.

Li, T., Chang, C.Y., Jin, D.Y., Lin, P.J., Khvorova, A., and Stafford, D.W. (2004). Identification of the gene for vitamin K epoxide reductase. Nature *427*, 541-544.

Limdi, N.A., and Veenstra, D.L. (2008). Warfarin Pharmacogenetics. Pharmacotherapy *28*, 1084–1097.

Liska, D.J. (1998). The detoxification enzyme systems. Altern. Med. Rev. *3*, 187-198.

Lu, H.F., Shih, M.C., Hsueh, S.C., Chen, C.M., Chang, J.Y., and Chang, J.G. (2005). Molecular analysis of the thiopurine S-methyltransferase alleles in Bolivians and Tibetans. J. Clin. Pharm. Ther. *30*, 491-496.

Lv, X., Tang, S., Xia, Y., Zhang, Y., Wu, S., Yang, Z., Li, X., Tu, D., Chen, Y., Deng, P., *et al.* (2012). *NAT2* genetic polymorphisms and anti-tuberculosis drug-induced hepatotoxicity in Chinese community population. Ann. Hepatol. *11*, 700-707.

Ma, M.K., Woo, M.H., and McLeod, H.L. (2002). Genetic basis of drug metabolism. Am. J. Health Syst. Pharm. *59*, 2061-2069.

Markatos, C.N., Grouzi, E., Politou, M., Gialeraki, A., Merkouri, E., Panagou, I., Spiliotopoulou, I., and Travlou, A. (2008). *VKORC1* and *CYP2C9* allelic variants influence acenocoumarol dose requirements in Greek patients. Pharmacogenomics *9*, 1631-1638.

Mas, S., Crescenti, A., Vidal-Taboada, J.M., Bergoñon, S., Cuevillas, F., Laso, N., Molina, R., Ballesta, A., and Lafuente, A. (2005). Simultaneous genotyping of *CYP2C9*2*, **3*, and 5′ flanking region (C-1189T) polymorphisms in a Spanish population through a new minisequencing multiplex single-base extension analysis. Eur. J. Clin. Pharmacol. *61*, 635-641.

McLeod, H.L., Krynetski, E.Y., Relling, M.V., and Evans, W.E. (2000). Genetic polymorphism of thiopurine methyltransferase and its clinical relevance for childhood acute lymphoblastic leukemia. Leukemia *14*, 567-572.

Mendizabal, I., Valente, C., Gusmão, A., Alves, C., Gomes, V., Goios, A., Parson, W., Calafell, F., Alvarez, L., Amorim, A.*, et al.* (2011). Reconstructing the Indian Origin and Dispersal of the European Roma: A Maternal Genetic Perspective. PLoS One *6*, e15988.

Miao, L., Yang, J., Huang, C., and Shen, Z. (2007). Contribution of age, body weight, and *CYP2C9* and *VKORC1* genotype to the anticoagulant response to warfarin: proposal for a new dosing regimen in Chinese patients. Eur. J. Clin. Pharmacol. *63*, 1135-1141.

Moaddeb, J., and Haga, S.B. (2013). Pharmacogenetic testing: current evidence of clinical utility. Ther. Adv. Drug Saf. *4*, 155-169.

Montes, R., Ruiz de Gaona, E., Martínez-González, M.Á., Alberca, I., and Hermida, J. (2006). The *c.−1639G > A* polymorphism of the *VKORC1* gene is a major determinant of the response to acenocoumarol in anticoagulated patients. Br. J. Haematol. *133*, 183-187.

Mushiroda, T., Ohnishi, Y., Saito, S., Takahashi, A., Kikuchi, Y., Saito, S., Shimomura, H., Wanibuchi, Y., Suzuki, T., Kamatani, N.*, et al.* (2006). Association of *VKORC1* and *CYP2C9* polymorphisms with warfarin dose requirements in Japanese patients. J. Hum. Genet. *51*, 249-253.

Musshoff, F., Stamer, U.M., and Madea, B. (2010). Pharmacogenetics and forensic toxicology. Forensic Sci. Int. *203*, 53-62.

Nahar, R., Deb, R., Saxena, R., Puri, R.D., and Verma, I.C. (2013). Variability in *CYP2C9* allele frequency: a pilot study of its predicted impact on warfarin response among healthy South and North Indians. Pharmacol. Rep. *65*, 187-194.

Olinga, P., Elferink, M.G.L., Draaisma, A.L., Merema, M.T., Castell, J.V., Pérez, G., and Groothuis, G.M.M. (2008). Coordinated induction of drug transporters and phase I and II metabolism in human liver slices. Eur. J. Pharm. Sci. *33*, 380-389.

Oliveira, E., Marsh, S., van Booven, D.J., Amorim, A., Prata, M.J., and McLeod, H.L. (2007a). Pharmacogenetically relevant polymorphisms in Portugal. Pharmacogenomics *8*, 703-712.

Oliveira, E., Quental, S., Alves, S., Amorim, A., and Prata, M.J. (2007b). Do the distribution patterns of polymorphisms at the thiopurine S-methyltransferase locus in sub-Saharan populations need revision? Hints from Cabinda and Mozambique. Eur. J. Clin. Pharmacol. *63*, 703-706.

Palovaara, S., Tybring, G., and Laine, K. (2003). The effect of ethinyloestradiol and levonorgestrel on the *CYP2C19*-mediated metabolism of omeprazole in healthy female subjects. Br. J. Clin. Pharmacol. *56*, 232-237.

Pavlovic, S. (2009). *TPMT* gene polymorphisms: on the doorstep of personalized medicine. Indian J. Med. Res. *129*, 478-480.

Pedersen, R.S., Brasch-Andersen, C., Sim, S.C., Bergmann, T.K., Halling, J., Petersen, M.S., Weihe, P., Edvardsen, H., Kristensen, V.N., Brøsen, K.*, et al.* (2010). Linkage disequilibrium between the *CYP2C19*17* allele and wildtype *CYP2C8* and *CYP2C9* alleles: identification of *CYP2C* haplotypes in healthy Nordic populations. Eur. J. Clin. Pharmacol. *66*, 1199-1205.

Pelotti, S., and Bini, C. (2011). Forensic Pharmacogenetics. In Forensic Medicine - From Old Problems to New Challenges, D.N. Vieira, ed. (http://www.intechopen.com/books/forensic-medicine-from-old-problems-to-new-challenges/forensic-pharmacogenetics: InTech).

Pereira, V., Gusmão, L., Valente, C., Pereira, R., Carneiro, J., Gomes, I., Morling, N., Amorim, A., and João Prata, M. (2012). Refining the genetic portrait of Portuguese Roma through X-chromosomal markers. Am. J. Phys. Anthropol. *148*, 389-394.

Pilgrim, J.L., Gerostamoulos, D., and Drummer, O.H. (2011). Review: Pharmacogenetic aspects of the effect of cytochrome P450 polymorphisms on serotonergic drug metabolism, response, interactions, and adverse effects. Forensic Sci. Med. Pathol. *7*, 162-184.

Poolsup, N., Li Wan Po, A., and Knight, T.L. (2000). Pharmacogenetics and psychopharmacotherapy. J. Clin. Pharm. Ther. *25*, 197-220.

Puehringer, H., Loreth, R.M., Klose, G., Schreyer, B., Krugluger, W., Schneider, B., and Oberkanins, C. (2010). *VKORC1* −1639G>A and *CYP2C9*3* are the major genetic predictors of phenprocoumon dose requirement. Eur. J. Clin. Pharmacol. *66*, 591-598.

Qu, L., Li, X., Wu, G., and Yang, N. (2005). Efficient and sensitive method of DNA silver staining in polyacrylamide gels. Electrophoresis *26*, 99-101.

Quental, S., Gusmão, A., Rodríguez-Pombo, P., Ugarte, M., Vilarinho, L., Amorim, A., and Prata, M.J. (2009). Revisiting MSUD in Portuguese Gypsies: Evidence for a Founder Mutation and for a Mutational Hotspot within the *BCKDHA* Gene. Ann Hum Genet *73*, 298-303.

Rai, A.J., Udar, N., Saad, R., and Fleisher, M. (2009). A Multiplex Assay for Detecting Genetic Variations in *CYP2C9*, *VKORC1*, and *GGCX* Involved in Warfarin Metabolism. Clin. Chem. *55*, 823-826.

Ramsjö, M., Aklillu, E., Bohman, L., Ingelman-Sundberg, M., Roh, H.K., and Bertilsson, L. (2010). CYP2C19 activity comparison between Swedes and Koreans: effect of genotype, sex, oral contraceptive use, and smoking. Eur. J. Clin. Pharmacol. *66*, 871-877.

Rieder, M.J., Reiner, A.P., Gage, B.F., Nickerson, D.A., Eby, C.S., McLeod, H.L., Blough, D.K., Thummel, K.E., Veenstra, D.L., and Rettie, A.E. (2005). Effect of *VKORC1* Haplotypes on Transcriptional Regulation and Warfarin Dose. N. Engl. J. Med. *352*, 2285-2293.

Rios-Santos, F., and Magno, L.A. (2012). Pharmacogenetics and Metabolism: Past, Present and Future. In Topics on Drug Metabolism, D.J. Paxton, ed. (http://www.intechopen.com/books/topics-on-drug-metabolism/pharmacogenetics-and-metabolism-past-present-and-future-: InTech).

Roberts, R.L., and Barclay, M.L. (2012). Current relevance of pharmacogenetics in immunomodulation treatment for Crohn's disease. J. Gastroenterol. Hepatol. *27*, 1546-1554.

Román, M., Cabaleiro, T., Ochoa, D., Novalbos, J., Chaparro, M., Gisbert, J.P., and Abad-Santos, F. (2012). Validation of a Genotyping Method for Analysis of *TPMT* Polymorphisms. Clin. Ther. *34*, 878-884.

Rossi, A.M., Bianchi, M., Guarnieri, C., Barale, R., and Pacifici, G.M. (2001). Genotype–phenotype correlation for thiopurine S-methyltransferase in healthy Italian subjects. Eur. J. Clin. Pharmacol. *57*, 51-54.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. *132*, 365-386.

Rushmore, T.H., and Kong, A.N. (2002). Pharmacogenomics, regulation and signaling pathways of phase I and II drug metabolizing enzymes. Curr. Drug Metab. *3*, 481-490.

Saad, Y.B., and Langaee, T. (2011). Genetic Polymorphisms of *CYP2C9*: Comparison of Prevalence in the Lebanese Population with Other Populations. Pharmacology & Pharmacy *2*, 88-93.

Sajantila, A., Palo, J.U., Ojanperä, I., Davis, C., and Budowle, B. (2010). Pharmacogenetics in medico-legal context. Forensic Sci. Int. *203*, 44-52.

Saminathan, R., Bai, J., Sadrolodabaee, L., Karthik, G.M., Singh, O., Subramaniyan, K., Ching, C.B., Chen, W.N., and Chowbay, B. (2010). *VKORC1* Pharmacogenetics and Pharmacoproteomics in Patients on Warfarin Anticoagulant Therapy: Transthyretin Precursor as a Potential Biomarker. PloS One *5*, e15064.

Samochatova, E.V., Chupova, N.V., Rudneva, A., Makarova, O., Nasedkina, T.V., Fedorova, O.E., Glotov, A.S., Kozhekbaeva, Z., Maiorova, O.A., Roumyantsev, A.G.*, et al.* (2009). *TPMT* genetic variations in populations of the Russian Federation. Pediatr. Blood Cancer *52*, 203-208.

Santos, P., Soares, R., Santos, D., Nascimento, R., Coelho, G., Nicolau, J., Mill, J., Krieger, J., and Pereira, A. (2011). *CYP2C19* and *ABCB1* gene polymorphisms are differently distributed according to ethnicity in the Brazilian general population. BMC Med. Genet. *12*, 13.

Sayitoglu, M.A., Yildiz, I., Hatirnaz, O., and Ozbek, U. (2006). Common cytochrome p4503A (*CYP3A4* and *CYP3A5*) and thiopurine S-methyl transferase (*TPMT*) polymorphisms in Turkish population. Turk. J. Med. Sci. *36*, 11-15.

Scibona, P., Redal, M.A., Garfi, L.G., Arbelbide, J., Argibay, P.F., and Belloso, W.H. (2012). Prevalence of *CYP2C9* and *VKORC1* alleles in the Argentine population and implications for prescribing dosages of anticoagulants. Genet. Mol. Res. *11*, 70-76.

Scordo, M.G., Aklillu, E., Yasar, U., Dahl, M.L., Spina, E., and Ingelman-Sundberg, M. (2001). Genetic polymorphism of cytochrome P450 2C9 in a Caucasian and a black African population. Br. J. Clin. Pharmacol. *52*, 447-450.

Scordo, M.G., Caputi, A.P., D'Arrigo, C., Fava, G., and Spina, E. (2004). Allele and genotype frequencies of *CYP2C9*, *CYP2C19* and *CYP2D6* in an Italian population. Pharmacol. Res. *50*, 195-200.

Scott, S.A. (2011). Personalizing medicine with clinical pharmacogenetics. Genet. Med. *13*, 987-995.

Scott, S.A., Edelmann, L., Kornreich, R., and Desnick, R.J. (2008). Warfarin Pharmacogenetics: *CYP2C9* and *VKORC1* Genotypes Predict Different Sensitivity and Resistance Frequencies in the Ashkenazi and Sephardi Jewish Populations. Am. J. Hum. Genet. *82*, 495-500.

Scott, S.A., Khasawneh, R., Peter, I., Kornreich, R., and Desnick, R.J. (2010). Combined *CYP2C9*, *VKORC1* and *CYP4F2* frequencies among racial and ethnic groups. Pharmacogenomics *11*, 781-791.

Selinski, S., Blaszkewicz, M., Lehmann, M.L., Ovsiannikov, D., Moormann, O., Guballa, C., Kress, A., Truss, M.C., Gerullis, H., Otto, T.*, et al.* (2011). Genotyping *NAT2* with only two SNPs (rs1041983 and rs1801280) outperforms the tagging SNP rs1495741 and is equivalent to the conventional 7-SNP *NAT2* genotype. Pharmacogenet Genomics *21*, 673-678.

Shirai, N., Furuta, T., Moriyama, Y., Okochi, H., Kobayashi, K., Takashima, M., Xiao, F., Kosuge, K., Nakagawa, K., Hanai, H.*, et al.* (2001). Effects of *CYP2C19* genotypic differences in the metabolism of omeprazole and rabeprazole on intragastric pH. Aliment. Pharmacol. Ther. *15*, 1929-1937.

Silan, C., Dogan, O.T., Silan, F., Kukulguven, F.M., Asgun, H.F., Ozdemir, S., Uludag, A., Atik, S., Gungor, B., Akdur, S.*, et al.* (2012). The prevalence of *VKORC1* 1639 G>A and *CYP2C9*2*3* genotypes in patients that requiring anticoagulant therapy in Turkish population. Mol. Biol. Rep. *39*, 11017-11022.

Sipeky, C., Csongei, V., Jaromi, L., Safrany, E., Polgar, N., Lakner, L., Szabo, M., Takacs, I., and Melegh, B. (2009a). Vitamin K epoxide reductase complex 1 (*VKORC1*) haplotypes in healthy Hungarian and Roma population samples. Pharmacogenomics *10*, 1025-1032.

Sipeky, C., Lakner, L., Szabo, M., Takacs, I., Tamasi, V., Polgar, N., Falus, A., and Melegh, B. (2009b). Interethnic differences of *CYP2C9* alleles in healthy Hungarian and Roma population samples: Relationship to worldwide allelic frequencies. Blood Cells Mol. Dis. *43*, 239-242.

Sipeky, C., Weber, A., Szabo, M., Melegh, B., Janicsek, I., Tarlos, G., Szabo, I., Sumegi, K., and Melegh, B. (2013). High prevalence of *CYP2C19*2* allele in Roma samples: study on Roma and Hungarian population samples with review of the literature. Mol. Biol. Rep. *40*, 4727-4735.

Smires, F.Z., Moreau, C., Habbal, R., Siguret, V., Fadili, S., Golmard, J.L., Assaidi, A., Beaune, P., Loriot, M.A., and Nadifi, S. (2012). Influence of genetics and non-genetic factors on acenocoumarol maintenance dose requirement in Moroccan patients. J. Clin. Pharm. Ther. *37*, 594-598.

Smith, M.A., Marinaki, A.M., Arenas, M., Shobowale-Bakre, M., Lewis, C.M., Ansari, A., Duley, J., and Sanderson, J.D. (2009). Novel pharmacogenetic markers for treatment outcome in azathioprine-treated inflammatory bowel disease. Aliment. Pharmacol. Ther. *30*, 375-384.

Spear, B.B., Heath-Chiozzi, M., and Huff, J. (2001). Clinical application of pharmacogenetics. Trends Mol. Med. *7*, 201-204.

Spire-Vayron de la Moureyre, C., Debuysere, H., Mastain, B., Vinner, E., Marez, D., Lo Guidice, J.M., Chevalier, D., Brique, S., Motte, K., Colombel, J.F.*, et al.* (1998). Genotypic and phenotypic analysis of the polymorphic thiopurine S-methyltransferase gene (*TPMT*) in a European population. Br. J. Pharmacol. *125*, 879-887.

Stachulski, A., and Lennard, M. (2000). Drug Metabolism: The Body's Defense against Chemical Attack. J. Chem. Educ. *77*, 349-353.

Suarez-Kurtz, G., Sortica, V.A., Vargens, D.D., Bruxel, E.M., Petz-Erler, M.-L., Tsuneto, L.T., and Hutz, M.H. (2012). Impact of population diversity on the prediction of 7-SNP *NAT2* phenotypes using the tagSNP rs1495741 or paired SNPs. Pharmacogenet. Genomics *22*, 305-309.

Sugimoto, K., Uno, T., Yamazaki, H., and Tateishi, T. (2008). Limited frequency of the *CYP2C19*17* allele and its minor role in a Japanese population. Br. J. Clin. Pharmacol. *65*, 437-439.

Suriapranata, I., Tjong, W., Wang, T., Utama, A., Raharjo, S., Yuniadi, Y., and Tai, S. (2011). Genetic factors associated with patient-specific warfarin dose in ethnic Indonesians. BMC Med. Genet. *12*, 80.

Swen, J.J., and Guchelaar, H.J. (2012). Just how feasible is pharmacogenetic testing in the primary healthcare setting? Pharmacogenomics *13*, 507-509.

Swen, J.J., Huizinga, T.W., Gelderblom, H., de Vries, E.G., Assendelft, W.J., Kirchheiner, J., and Guchelaar, H.J. (2007). Translating Pharmacogenomics: Challenges on the Road to the Clinic. PLoS Med. *4*, e209.

Taniguchi, C., and Guengerich, F.P. (2012). Drug Metabolism. In Principles of Pharmacology: The Pathophysiologic Basis of Drug Therapy, D.E. Golan, A.H. Tashijian, E.J. Armstrong, and A.W. Armstrong, eds. (Philadelphia Wolters Kluwer Health/Lippincott Williams & Wilkins), pp. 43-54.

Teixeira, R., Lopes, M., and Santos, P. (2013). Tuberculosis Pharmacogenetics: State of The Art. In Tuberculosis - Current Issues in Diagnosis and Management, D.B. Mahboub, ed. (http://www.intechopen.com/books/tuberculosis-current-issues-in-diagnosis-and-management/tuberculosis-pharmacogenetics-state-of-the-art: InTech).

Teixeira, R.L., Morato, R.G., Cabello, P.H., Muniz, L.M., Moreira Ada, S., Kritski, A.L., Mello, F.C., Suffys, P.N., Miranda, A.B., and Santos, A.R. (2011). Genetic polymorphisms of NAT2, CYP2E1 and GST enzymes and the occurrence of antituberculosis drug-induced hepatitis in Brazilian TB patients. Mem. Inst. Oswaldo Cruz *106*, 716-724.

Tiwari, A., Souza, R., and Müller, D. (2009). Pharmacogenetics of anxiolytic drugs. J. Neural Transm. *116*, 667-677.

Toft, N., Nygaard, U., Gregers, J., and Schmiegelow, K. (2006). Genetic analyses of thiopurine methyltransferase polymorphisms in Greenlandic and Danish populations. Acta Paediatr. *95*, 1665-1667.

Umamaheswaran, G., Krishna Kumar, D., Kayathiri, D., Rajan, S., Shewade, D.G., Dkhar, S.A., Manjunath, S., Ushakiran, P., Reneega, G., Ritushree, K.*, et al.* (2012). Inter and intra-ethnic differences in the distribution of the molecular variants of *TPMT, UGT1A1* and *MDR1* genes in the South Indian population. Mol. Biol. Rep. *39*, 6343-6351.

Uno, T., Niioka, T., Hayakari, M., Yasui-Furukori, N., Sugawara, K., and Tateishi, T. (2007). Absolute bioavailability and metabolism of omeprazole in relation to *CYP2C19* genotypes following single intravenous and oral administrations. Eur. J. Clin. Pharmacol. *63*, 143-149.

Valente, C., Gomes, I., Pereira, V., Amorim, A., Gusmão, L., and Prata, M.J. (2009). Association between STRs from the X chromosome in a sample of Portuguese Gypsies. Forensic Sci. Int. *2*, 391-393.

Vallone, P.M., and Butler, J.M. (2004). AutoDimer: a screening tool for primer-dimer and hairpin structures. Biotechniques *37*, 226-231.

Voora, D., and Ginsburg, G.S. (2012). Clinical Application of Cardiovascular Pharmacogenetics. J. Am. Coll. Cardiol. *60*, 9-20.

Vorob'eva, N.M., Panchenko, E.P., Dobrovol'skiĭ, A.B., Titaeva, E.V., Khasanova, Z.B., Konovalova, N.V., Postnov, A., and Kirienko, A. (2011). Polymorphisms of genes *CYP2C9* and *VKORC1* in patients with venous thromboembolic complications in Moscow population: effects on stability of anticoagulant therapy and frequency of hemorrhage. Ter. Arkh. *83*, 59-65.

Wadelius, M., Chen, L.Y., Lindh, J.D., Eriksson, N., Ghori, M.J.R., Bumpstead, S., Holm, L., McGinnis, R., Rane, A., and Deloukas, P. (2009). The largest prospective warfarin-treated cohort supports genetic forecasting. Blood *113*, 784-792.

Wang, D., Chen, H., Momary, K.M., Cavallari, L.H., Johnson, J.A., and Sadée, W. (2008). Regulatory polymorphism in vitamin K epoxide reductase complex subunit 1 (*VKORC1*) affects gene expression and warfarin dose requirement. Blood *112*, 1013-1021.

Wang, L., Pelleymounter. L., Weinshilboum, R., Johnson, J.A., Hebert, J.M., Altman, R.B., and Klein, T.E. (2010). Very important pharmacogene summary: thiopurine S-methyltransferase. Pharmacogenet. Genomics *20*, 401–405.

Wang, P.Y., Xie, S.Y., Hao, Q., Zhang, C., and Jiang, B.F. (2012). *NAT2* polymorphisms and susceptibility to anti-tuberculosis drug-induced liver injury: a meta-analysis. Int. J. Tuberc. Lung Dis. *16*, 589-595.

Weinshilboum, R., and Wang, L. (2004). Pharmacogenomics: bench to bedside. Nat. Rev. Drug Discov. *3*, 739-748.

Weinshilboum, R., and Wang, L. (2006). Pharmacogenetics and Pharmacogenomics: Development, Science, and Translation. Annu. Rev. Genomics Hum. Genet. *7*, 223-245.

Woelderink, A., Ibarreta, D., Hopkins, M.M., and Rodriguez-Cerezo, E. (2005). The current clinical practice of pharmacogenetic testing in Europe: *TPMT* and *HER2* as case studies. Pharmacogenomics J. *6*, 3-7.

Wong, S.H., Wagner, M.A., Jentzen, J.M., Schur, C., Bjerke, J., Gock, S.B., and Chang, C.C. (2003). Pharmacogenomics as an aspect of molecular autopsy for forensic pathology/toxicology: does genotyping *CYP 2D6* serve as an adjunct for certifying methadone toxicity? J. Forensic Sci. *48*, 1406-1415.

Wu, A.H. (2007). Use of genetic and nongenetic factors in warfarin dosing algorithms. Pharmacogenomics *8*, 851-861.

Xiao, Z.S., Goldstein, J.A., Xie, H.G., Blaisdell, J., Wang, W., Jiang, C.H., Yan, F.X., He, N., Huang, S.L., Xu, Z.H*., et al.* (1997). Differences in the Incidence of the *CYP2C19* Polymorphism Affecting the S-Mephenytoin Phenotype in Chinese Han and Bai Populations and Identification of a New Rare *CYP2C19* Mutant Allele. J. Pharmacol. Exp. Ther. *281*, 604-609.

Xie, H.G., Prasad, H.C., Kim, R.B., and Stein, C.M. (2002). *CYP2C9* allelic variants: ethnic distribution and functional significance. Adv. Drug Deliv. Rev. *54*, 1257-1270.

Xu, C., Li, C., and Kong, A.N. (2005). Induction of phase I, II and III drug metabolism/transport by xenobiotics. Arch. Pharm. Res. *28*, 249-268.

Yamada, S., Onda, M., Kato, S., Matsuda, N., Matsuhisa, T., Yamada, N., Miki, M., and Matsukura, N. (2001). Genetic differences in *CYP2C19* single nucleotide polymorphisms among four Asian populations. J. Gastroenterol. *36*, 669-672.

Yang, J.Q., Morin, S., Verstuyft, C., Fan, L.A., Zhang, Y., Xu, C.D., Barbu, V., Jaillon, P., and Becquemont, L. (2003). Frequency of cytochrome P450 2C9 allelic variants in the Chinese and French populations. Fundam. Clin. Pharmacol. *17*, 373-376.

Yang, Y.S., Wong, L.P., Lee, T.C., Mustafa, A.M., Mohamed, Z., and Lang, C.C. (2004). Genetic polymorphism of cytochrome P450 2C19 in healthy Malaysian subjects. Br. J. Clin. Pharmacol. *58*, 332-335.

Yasui-Furukori, N., Takahata, T., Nakagami, T., Yoshiya, G., Inoue, Y., Kaneko, S., and Tateishi, T. (2004). Different inhibitory effect of fluvoxamine on omeprazole metabolism between *CYP2C19* genotypes. Br. J. Clin. Pharmacol. *57*, 487-494.

Yin, T., and Miyata, T. (2007). Warfarin dose and the pharmacogenomics of *CYP2C9* and *VKORC1* — Rationale and perspectives. Thromb. Res. *120*, 1-10.

Yoshizawa, M., Hayashi, H., Tashiro, Y., Sakawa, S., Moriwaki, H., Akimoto, T., Doi, O., Kimura, M., Kawarasaki, Y., Inoue, K.*, et al.* (2009). Effect of *VKORC1*-1639 G>A polymorphism, body weight, age, and serum albumin alterations on warfarin response in Japanese patients. Thromb. Res. *124*, 161-166.

Yousef, A.M., Bulatova, N.R., Newman, W., Hakooz, N., Ismail, S., Qusa, H., Zahran, F., Anwar Ababneh, N., Hasan, F., Zaloom, I.*, et al.* (2012). Allele and genotype frequencies of the polymorphic cytochrome P450 genes (*CYP1A1*, *CYP3A4*, *CYP3A5*, *CYP2C9* and *CYP2C19*) in the Jordanian population. Mol. Biol. Rep. *39*, 9423-9433.

Zabala-Fernández, W., Barreiro-de Acosta, M., Echarri, A., Carpio, D., Lorenzo, A., Castro, J., Martínez-Ares, D., Pereira, S., Martin-Granizo, I., Corton, M.*, et al.* (2011). A pharmacogenetics study of *TPMT* and *ITPA* genes detects a relationship with side effects and clinical response in patients with inflammatory bowel disease receiving Azathioprine. J. Gastrointestin. Liver Dis. *20*, 247-253.

Zambon, C.F., Pengo, V., Padrini, R., Basso, D., Schiavon, S., Fogar, P., Nisi, A., Frigo, A.C., Moz, S., Pelloso, M.*, et al.* (2010). *VKORC1, CYP2C9* and *CYP4F2* genetic-based algorithm for warfarin dosing: an Italian retrospective study. Pharmacogenomics *12*, 15-25.

Zanger, U., Turpeinen, M., Klein, K., and Schwab, M. (2008). Functional pharmacogenetics/genomics of human cytochromes P450 involved in drug biotransformation. Anal. Bioanal. Chem. *392*, 1093-1108.

# VIII. ANNEXES

**Annex 1.** Supplementary data for the ten *loci* studied.

| | Portuguese non-Gypsies | | | Portuguese Gypsies | | |
|---|---|---|---|---|---|---|
| **Locus** | **Ho** | **He** | **P-value** | **Ho** | **He** | **P-value** |
| c.238G>C | | Monomorphic locus | | | Monomorphic locus | |
| c.460G>A | 0.05714 | 0.05591 | 1.00000±0.00000 | 0.00862 | 0.00862 | 1.00000±0.00000 |
| c.719A>G | 0.05744 | 0.05591 | 1.00000±0.00000 | 0.00862 | 0.00862 | 1.00000±0.00000 |
| c.664G>A | | Monomorphic locus | | 0.00862 | 0.00862 | 1.00000±0.00000 |
| c.430C>T | 0.31429 | 0.32230 | 1.00000±0.00000 | 0.18103 | 0.19313 | 0.61789±0.00050 |
| c.1075A>C | 0.11429 | 0.10853 | 1.00000±0.00000 | 0.15789 | 0.14607 | 1.00000±0.00000 |
| c.681G>A | 0.3000 | 0.27657 | 0.67627±0.00047 | 0.30172 | 0.33464 | 0.27913±0.00044 |
| c.282C>T | 0.45714 | 0.42302 | 0.57576±0.00050 | 0.42609 | 0.49474 | 0.18493±0.00041 |
| c.341T>C | 0.44286 | 0.50103 | 0.34618±0.00050 | 0.38261 | 0.43835 | 0.20270±0.00039 |
| c.-1639G>A | 0.48571 | 0.49702 | 1.00000±0.00000 | 0.47826 | 0.48380 | 1.00000±0.00000 |

He, Expected Heterozygosity

Ho, Observed Heterozygosity